DUKE MATHEMATICAL JOURNAL

VOLUME 4
1938

DUKE UNIVERSITY PRESS
DURHAM, N. C.

CONTENTS

VOLUME 4, 1938

ADER, OLIN B. An affine invariant of convex regions ... 291
AGNEW, RALPH PALMER. Linear functionals satisfying prescribed conditions ... 55
BATEMAN, H. Spheroidal and bipolar coördinates ... 39
BELL, P. O. Covariant configurations related to analytic curved surfaces ... 590
BOAS, R. P., JR. Some gap theorems for power series ... 176
BOAS, R. P., JR. Tauberian theorems for $(C, 1)$ summability ... 227
BOCHNER, S. Differentiable and Riemann metric ... 51
BOCHNER, S. Functions of integrable square in several complex variables ... 635
BOHR, HARALD, and FLANDERS, DONALD A. Algebraic functions of analytic almost periodic functions ... 779
CHANLER, J. H., and THRALL, R. M. Ternary trilinear forms in the field of complex numbers ... 678
DOOB, J. L. One-parameter families of transformations ... 752
DRIBIN, D. M. Prüfer ideals in commutative rings ... 737
DUTHIE, WILLIAM D. Boolean functions of bounded variation ... 600
FLANDERS, DONALD A., and BOHR, HARALD. Algebraic functions of analytic almost periodic functions ... 779
FORT, TOMLINSON. The summability of exponential and factorial series ... 793
GARABEDIAN, H. L., and RANDELS, W. C. Theorems on Riesz means ... 529
GENTRY, FRANK C. Quaternary Cremona groups of ternary type ... 107
GOLDSTINE, H. H. Weakly complete Banach spaces ... 125
HALL, D. W., and SCHWEIGERT, G. E. Properties of invariant sets under pointwise periodic homeomorphisms ... 719
HALL, MARSHALL. Equidistribution of residues in sequences ... 691
HALPERIN, I., and PITT, H. R. Integral inequalities connected with differential operators ... 613
HAMILTON, HUGH J. Change of dimension in sequence transformations ... 341
HAMILTON, HUGH J. A generalization of multiple sequence transformations ... 343
HOPKINS, CHARLES. The half-group of cosets belonging to a group ... 30
HOPKINS, CHARLES. Nil-rings with minimal condition for admissible left ideals ... 664
HSÜ, HSIEN-YÜ. Certain integrals and infinite series involving ultraspherical polynomials and Bessel functions ... 374
JACKSON, DUNHAM. Orthogonal polynomials in three variables ... 441
JACOBSON, N. Simple Lie algebras over a field of characteristic zero ... 534
JOHN, FRITZ. The ultrahyperbolic differential equation with four independent variables ... 300

KASNER, EDWARD. Characterization of the conformal group and the equilong group by horn angles 95
KETCHUM, P. W. Infinite systems of linear equations and expansions of analytic functions 668
KRALL, H. L. Certain differential equations for Tchebycheff polynomials. 705
LEHMER, D. H. A cotangent analogue of continued fractions 323
LEIGHTON, WALTER. Sufficient conditions for the convergence of a continued fraction 775
LESTER, CAROLINE A. A determination of the automorphisms of certain algebraic fields 277
LEVINSON, NORMAN. Integral functions bounded on sequences of points.. 170
McCOY, NEAL H. Subrings of infinite direct sums 486
McSHANE, E. J. Some existence theorems for problems in the calculus of variations 132
MAC LANE, SAUNDERS. A lattice formulation for transcendence degrees and p-bases 455
MACQUEEN, M. L. The projections of the asymptotic curves 412
MARCINKIEWICZ, J., and ZYGMUND, A. Proof of a gap theorem 469
MARCINKIEWICZ, J., and ZYGMUND, A. A theorem of Lusin 473
MARTIN, W. T., and WIENER, N. Taylor's series of functions of smooth growth in the unit circle 384
MATISON, HARRY. Certain integral functions related to exponential sums. 9
MONTGOMERY, DEANE, and ZIPPIN, LEO. Compact Abelian transformation groups 363
MORSE, MARSTON. The index theorem in the calculus of variations 231
ORE, OYSTEIN. Structures and group theory. II 247
PALL, GORDON. On the factorization of generalized quaternions 696
PETTIS, B. J. Linear functionals and completely additive set functions ... 552
PIERCE, JESSE. Solutions of a differential equation of the first order and first degree in the vicinity of branch points of the solution 650
PIERCE, JESSE. Solutions of systems of differential equations in infinitely many unknowns by infinite series of definite integrals 566
PITT, H. R. A remark on Wiener's general Tauberian theorem 437
PITT, H. R., and HALPERIN, I. Integral inequalities connected with differential operators 613
PITT, H. R., and WIENER, N. On absolutely convergent Fourier-Stieltjes transforms 420
RADÓ, TIBOR. On absolutely continuous transformations in the plane 189
RANDELS, W. C., and GARABEDIAN, H. L. Theorems on Riesz means 529
RINEHART, R. F. Commutative algebras which are polynomial algebras.. 725
RISSELMAN, W. C. Approximation to the solution of a normal system of ordinary linear differential equations 640
SAKS, S. Integration in abstract metric spaces 408

Schweigert, G. E., and Hall, D. W. Properties of invariant sets under pointwise periodic homeomorphisms 719
Serbin, H. Factorization in principal ideal rings 656
Sewell, W. E. Degree of approximation by polynomials in z and $1/z$ 393
Simond, R. G. Relations between certain continuous transformations of sets 575
Smith, A. H. Summability of conjugate Fourier series 270
Szász, Otto. The jump of a function determined by its Fourier coefficients 401
Thrall, R. M., and Chanler, J. H. Ternary trilinear forms in the field of complex numbers 678
Tornheim, Leonard. Sums of n-th powers in fields of prime characteristic 359
Tuller, Annita. The measure of transitive geodesics on certain three-dimensional manifolds 78
Wehausen, John V. Transformations in linear topological spaces 157
Whitney, Hassler. Cross-sections of curves in 3-space 222
Whitney, Hassler. Tensor products of Abelian groups 495
Whyburn, G. T. Interior surface transformations 626
Whyburn, G. T. Interior transformations on certain curves 607
Whyburn, G. T. The mapping of Betti groups under interior transformations 1
Wiener, N., and Martin, W. T. Taylor's series of functions of smooth growth in the unit circle 384
Wiener, N., and Pitt, H. R. On absolutely convergent Fourier-Stieltjes transforms 420
Zippin, Leo, and Montgomery, Deane. Compact Abelian transformation groups 363
Zorn, Max. Discontinuous groups and allied topics, III: On a lemma about matrices 788
Zygmund, A., and Marcinkiewicz, J. Proof of a gap theorem 469
Zygmund, A., and Marcinkiewicz, J. A theorem of Lusin 473

THE MAPPING OF BETTI GROUPS UNDER INTERIOR TRANSFORMATIONS

By G. T. Whyburn

1. In this paper results will be established from which it follows that the one-dimensional rational Betti group of a compact metric set A under any interior transformation[1] on A maps homomorphically onto the corresponding group of the image of A provided the one-dimensional Betti number $p^1(A)$ is finite. Thus in any case the one-dimensional Betti number of a compact set is not increased[2] when the set undergoes an interior transformation. This presents, in part, a solution to a problem proposed by Eilenberg.[3] However, by making use of what are termed *nodal subsets* of a compact set, we are able to obtain results of considerably increased generality and precision.

We begin with an elementary characterization of interior transformations.

2. Theorem. *Let A and B be compact and $T(A) = B$ be continuous. In order that T be an interior transformation, it is necessary and sufficient that for any $\epsilon > 0$ there exist a $\delta > 0$ such that for any $y \in B$ and any $x \in T^{-1}(y)$, $T[V_\epsilon(x)] \supset V_\delta(y)$.*[4]

Proof. To prove the sufficiency, let U be any open set in A, let $T(U) = V$. Let $y \in V$, $x \in U \cdot T^{-1}(y)$ and let $\epsilon > 0$ be chosen so that $V_\epsilon(x) \subset U$. Then by hypothesis $T[V_\epsilon(x)] \supset V_\delta(y)$ for some $\delta > 0$. This gives

$$V = T(U) \supset T[V_\epsilon(x)] \supset V_\delta(y)$$

and hence V is open in B.

To establish the necessity of the condition, suppose on the contrary that for some $\epsilon > 0$ there exist a sequence y_i in B, $y_i \to y \in B$, a sequence x_i in A, $x_i \to x \in T^{-1}(y)$, $x_i \in T^{-1}(y_i)$, and a sequence of numbers $0 < \delta_i \to 0$ such that for no i does $T[V_\epsilon(x_i)] \supset V_{\delta_i}(y_i)$. But now $T[V_{\frac{1}{2}\epsilon}(x)] = V$ is open and contains

Received October 30, 1937.

[1] A continuous transformation $T(A) = B$ is *interior* provided every open set in A maps into an open set in B. See Stoïlow, Annales Scientifiques de l'Ecole Normale Supérieure, vol. 63 (1928), pp. 347–382; we follow the usual custom of omitting Stoïlow's second condition that no continuum in A map into a single point in B.

[2] For the case of this result where A is a graph or a 1-dimensional locally connected continuum, see my paper *Interior transformations on compact sets*, this Journal, vol. 3 (1937), pp. 370–381.

[3] Fundamenta Mathematicae, vol. 24 (1935), p. 175.

[4] We employ the notation $V_r(X)$ for the ϵ-neighborhood of the set X, i.e., the set of all points y such that $\rho(x, y) < r$ for some $x \in X$.

y and hence it must contain $V_\delta(y)$ for some $\delta > 0$; and there exists an i with $\delta_i < \frac{1}{2}\delta$, $x_i \subset V_{\frac{1}{2}\epsilon}(x)$, and $y_i \subset V_{\frac{1}{2}\delta}(y)$. This gives

$$V_\epsilon(x_i) \supset V_{\frac{1}{2}\epsilon}(x), \qquad V_{\delta_i}(y_i) \subset V_\delta(y).$$

Whence

$$T[V_\epsilon(x_i)] \supset T[V_{\frac{1}{2}\epsilon}(x)] \supset V_\delta(y) \supset V_{\delta_i}(y_i),$$

contrary to our supposition.

3. **Nodal subsets.** A closed subset N of a compact set M will be called a *nodal* subset of M provided the set $N \cdot \overline{M - N}$ contains at most one point.

It should be noted that the set $M - N$ may be vacuous, i.e., N may be equal to M. Also it may be remarked that, if N is a nodal subset of M, then $\overline{M - N}$ is also a nodal subset of M; and if M is a continuum and p is any cut point of M, then for any separation $M - p = M_1 + M_2$, $M_1 + p$ and $M_2 + p$ are nodal subsets of M. If M is a cyclic continuum, i.e., a continuum without cut points, then the only nodal subsets of M are M itself and the individual points of M.

(3.1) *If A is compact and $T(A) = B$ is interior, any nodal subset of A maps onto a nodal subset of B under T.*

For let N be a nodal subset of A. If $N \cdot \overline{A - N} = 0$, N is open in A; and hence $T(N)$ is open in B so that $T(N) \cdot \overline{B - T(N)} = 0$. Thus we may suppose $N \cdot \overline{A - N} = x \in A$. Since $N - x$ is open in A, it follows that $T(N) - T(x)$ is open in B. Accordingly $T(N) \cdot \overline{B - T(N)} \subset T(x)$, so that $T(N)$ is a nodal subset of B.

(3.2) *If $T(A) = B$ is interior, where B is a cyclic continuum, any non-degenerate nodal subset of A maps onto all of B under T.*

Clearly this follows at once from (3.1) since the only non-degenerate nodal subset of B is B itself.

4. THEOREM. *Let $T(A) = B$ be interior, where A is compact, let f be the homomorphism generated in the cycle group of A by T (rational coefficient field), let N be any nodal subset of A and let $M = T(N)$. For any $\epsilon > 0$ there exists a $\delta > 0$ such that if Z is any 1-dimensional rational δ-cycle*[5] *in M there exists a 1-dimensional rational ϵ-cycle C in N such that $f(C) \underset{\epsilon}{\sim} Z$.*

Proof. Let $K = \overline{A - N}$. It will be apparent that the proof given here will be only greatly simplified in case either $N = A$ or $N \cdot M = 0$. Hence we may (and do) assume that $N \cdot M = p \in A$. Let $q = T(p)$.

It is obvious from the proof given in §2 that the necessity part of the theorem which was proved there can be stated in the following slightly more general form which suits our purposes in the present proof.

[5] That is, the "edges" of the cycle are of diameter $<\delta$. It should be noted that our polygons and complexes are abstract—only their vertices are actually realized in the sets concerned.

(4.1) *If $A = A_1 + A_2$, where A_1 and A_2 are closed, then for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $y \epsilon T(A_1)$ and $x \epsilon T^{-1}(y)[A_1 - V_\epsilon(A_1 \cdot A_2)]$, $T[A_1 \cdot V_\epsilon(x)] \supset V_\delta(y) \cdot T(A_1)$.*

Now to prove our theorem we shall treat first the case where Z is a single simply oriented simple closed polygon. Let $\epsilon' = \frac{1}{3}\epsilon$ and let $d < \epsilon'$ be a number δ determined in (4.1) using ϵ' for ϵ, N for A_1 and K for A_2. Let $d' = \frac{1}{4}d$ and let e be chosen so that $e < d'$ and $\delta(T[V_e(p)]) < d'$ so that $T[V_e(p)] \subset V_{d'}[T(p)]$. Let δ be determined from (4.1) using e for ϵ, $N = A_1$, $K = A_2$.[6] We next proceed to prove

(4.2) *If $s = [p_0, p_1, p_2, \cdots, p_n]$ is a simply oriented simple closed polygon in $M - M \cdot V_{d'}(q)$ such that $\rho(p_i, p_{i+1}) < \delta$ $(i = 0, 1, \cdots, n-1)$, $\rho(p_n, p_0) < d$, then there exists a simply oriented simple closed ϵ-polygon σ in N and an integer k such that $f(\sigma) = ks$.*

Proof of (4.2). Let $q_0 \epsilon T^{-1}(p_0) \cdot N$. Since $\rho(p_0, p_1) < \delta$, it follows by (4.1) and the choice of δ that there exists a point $q_1 \epsilon N \cdot T^{-1}(p_1)$ with $\rho(q_1, q_0) < e$. Similarly there exist a $q_2 \epsilon N \cdot T^{-1}(p_2)$ with $\rho(q_2, q_1) < e$, a $q_3 \epsilon N \cdot T^{-1}(p_3)$ with $\rho(q_3, q_2) < e$ and so on to $q_n \epsilon N \cdot T^{-1}(p_n)$ with $\rho(q_n, q_{n-1}) < e$.

Now if q_n is not an element of $V_{\epsilon'}(p)$, then since $\rho(q_n, q_0) < d$ it follows from (4.1) that there exists a $q_0^1 \epsilon N \cdot T^{-1}(p_0)$ with $\rho(q_0^1, q_n^1) < \epsilon'$. Then, continuing just as above, we find $q_1^1 \epsilon N \cdot T^{-1}(p_1)$, $q_2^1 \epsilon N \cdot T^{-1}(p_2)$, $\cdots$, $q_n^1 \epsilon N \cdot T^{-1}(p_n)$ with $\rho(q_0^1, q_1^1) < e$, $\cdots$, $\rho(q_{n-1}^1, q_n^1) < e$. And again if q_n^1 is not an element of $V_{\epsilon'}(p)$, this process may be repeated, and so on. Now either

(i) for some (least) i we get $q_n^i \epsilon V_{\epsilon'}(p)$, in which case the process ends, or

(ii) on account of the compactness of A we must eventually reach, for the first time, a point $q_i^j (i \leqq n)$ such that $\rho(q_i^j, q_{i+1}^m) < \epsilon$ $(m < j, n + 1 = 0)$. In case (ii) clearly the edges

$$q_0^{m+1} q_1^{m+1}, q_1^{m+1} q_2^{m+1}, \cdots, q_{n-1}^{m+1} q_n^{m+1}, q_n^{m+1} q_0^{m+2},$$
$$\cdots, q_{i-1}^j q_i^j, q_i^j q_{i+1}^m, q_{i+1}^m q_{i+2}^m, \cdots, q_n^m q_0^{m+1}$$

fit together to form a simple closed ϵ-polygon σ which maps onto s exactly $j - m$ times. Hence, if we set $k = j - m$, we have $f(\sigma) = ks$ as was required.

Now in case (i) where $q_n^i \epsilon V_{\epsilon'}(p)$ we repeat the same process proceeding in reverse order starting from q_0. That is, if q_0 is not an element of $V_{\epsilon'}(p)$, there exists a $q_n^{-1} \epsilon N \cdot T^{-1}(p_n)$ with $\rho(q_n^{-1}, q_0) < \epsilon'$. There exist $q_{n-1}^{-1} \epsilon N \cdot T^{-1}(p_{n-1})$, $\cdots$, $q_0^{-1} \epsilon N \cdot T^{-1}(p_0)$ with $\rho(q_n^{-1}, q_{n-1}^{-1}) < e$, $\cdots$, $\rho(q_1^{-1}, q_0^{-1}) < e$; again if q_0^{-1} is not an element of $V_{\epsilon'}(p)$, we find $q_n^{-2} \epsilon N \cdot T^{-1}(p_n)$ with $\rho(q_n^{-2}, q_0^{-1}) < \epsilon'$ and proceed as before. Now if this continues indefinitely, i.e., if we always have q_0^{-j} is not an element of $V_{\epsilon'}(p)$, it is clear that we can get our required polygon σ in the same manner as above under (ii). Hence we may suppose that for

[6] Note the following relations between the various numbers just chosen:

$$\epsilon > \tfrac{1}{3}\epsilon = \epsilon' > d > \tfrac{1}{4}d = d' > e > \delta.$$

some (least) j we get $q_0^{-j} \epsilon V_{\epsilon'}(p)$. But now since also $q_n^i \epsilon V_{\epsilon'}(p)$ we have $\rho(q_n^i, q_0^{-j}) < 2\epsilon' < \epsilon$. Accordingly the edges

$$q_0^{-j}q_1^{-j}, \cdots, q_{n-1}^{-j}q_n^{-j}, q_n^{-j}q_0^{-j+1}, q_0^{-j+1}q_1^{-j+1}, \cdots, q_0^i q_1^i, \cdots, q_{n-1}^i q_n^i, q_n^i q_0^{-j}$$

fit together to form a simple closed ϵ-polygon σ which maps onto s exactly $j + i + 1$ times. Hence if we set $k = j + i + 1$ we have $f(\sigma) = ks$ as required. Accordingly (4.2) is proved.

Continuing with our proof for the case where Z is a single simply oriented simple closed polygon, we next prove

(4.3) *There exist a finite number of oriented simple closed polygons* $s_1, s_2, s_3, \cdots, s_k$ *in* $M - V_{d'}(q) \cdot M$ *satisfying the hypothesis in* (4.2) *and such that*

$$Z \underset{\epsilon}{\sim} s_1 + s_2 + \cdots + s_k .$$

To show this we note first that the vertices of Z which are in $M - V_{d'}(q)$ fall into a finite number of maximal chains of successive vertices:

$$X_1 = [p_0^1, p_1^1, \cdots, p_{n_1}^1],$$
$$X_2 = [p_0^2, p_1^2, \cdots, p_{n_2}^2],$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$X_k = [p_0^k, p_1^k, \cdots, p_{n_k}^k],$$

where, in general, p_j^i follows p_n^m if $i > m$ or if $i = m$ and $j > n$, where the points preceding p_0^i and following $p_{n_i}^i$ in Z in every case belong to $V_{d'}(q)$, and where we have omitted all such chains with fewer than 3 vertices, i.e., $n_i \geqq 2$ in all cases.

Now let $Y_1, Y_2, \cdots, Y_k$ be the complementary chains in Z, i.e., $p_{n_i}^i$ and p_0^{i+1} are the initial and final points, respectively, of Y_i and furthermore we have

$$\text{(i)} \qquad Z = X_1 + Y_1 + X_2 + Y_2 + \cdots + X_k + Y_k$$

and for each $i \leqq k$,

$$\text{(ii)} \qquad Y_i \subset V_{2d'}(q).$$

For each $i \leqq k$ let x_i denote the simplex $(p_{n_i}^i, p_0^i)$, oriented as indicated, and let s_i denote the oriented polygon $X_i + x_i = [p_0^i, \cdots, p_{n_i}^i, p_0^i]$. Then, since $p_0^i + p_{n_i}^i \subset V_{2d'}(q)$, it follows that for each i, s_i satisfies the conditions of (4.2). Now by (i) we have

$$\text{(iii)} \qquad \begin{aligned} Z &= (X_1 + x_1) + (Y_1 - x_1) + \cdots + (X_k + x_k) + (Y_k - x_k) \\ &= (s_1 + s_2 + \cdots + s_k) + [(Y_1 - x_1) + \cdots + (Y_k - x_k)]. \end{aligned}$$

Furthermore

$$W = -x_1 + Y_1 - x_2 + Y_2 - \cdots - x_n + Y_n$$

is an oriented simple polygon lying entirely in $V_{2d'}(q) \subset V_d(q)$. Accordingly, we have

$$W \underset{d}{\sim} 0,$$

and hence surely

(iv) $$W \underset{\epsilon}{\sim} 0,$$

since $d < \epsilon$. Now (iii) reduces to

$$Z = \sum_1^k s_i + W.$$

Whence by (iv)

$$Z - \sum_1^k s_i = W \underset{\epsilon}{\sim} 0$$

or

$$Z \underset{\epsilon}{\sim} \sum s_i$$

as required by (4.3).

Now to prove our theorem in case Z is a single simple closed polygon we first apply (4.3) obtaining $s_1, s_2, \cdots, s_k$ satisfying (4.3) so that

$$Z \underset{\epsilon}{\sim} \sum s_i .$$

Then by virtue of (4.2), for each i there exists a simply oriented simple closed ϵ-polygon σ_i in N and an integer k_i such that

$$f(\sigma_i) = k_i s_i \qquad \text{or} \qquad f(\sigma_1/k_i) = s_i .$$

Hence, if we set

$$C = \sum_1^k \sigma_i/k_i ,$$

we have $f(C) = \sum_1^k s_i \underset{\epsilon}{\sim} Z$ as required.

Now in the general case, we can first express Z in the form

$$Z = a_1 Z_1 + a_2 Z_2 + \cdots + a_n Z_n ,$$

where a_i is rational and Z_i is a single simply oriented simple closed polygon. Then, applying the case already treated, for each i we find a 1-dimensional rational ϵ-cycle C_i in N such that

$$f(C_i) \underset{\epsilon}{\sim} Z_i .$$

Hence, if we set

$$C = a_1 C_1 + a_2 C_2 + a_3 C_3 + \cdots + a_n C_n ,$$

we have

$$f(C) = \sum_1^n a_i f(C_i) \underset{\epsilon}{\sim} \sum_1^n a_i Z_i = Z,$$

as was to be proved.

The same notation will be retained in the next section.

5. Theorem. *If V is any 1-dimensional rational Vietoris cycle*[7] *(or true cycle)*[8] *in M, for any $\epsilon > 0$ there exists a Vietoris cycle (or true cycle) W in N such that*

$$f(W) \underset{\epsilon}{\sim} V.$$

Proof. The case of a true cycle follows immediately from §4. Hence we suppose V is a Vietoris cycle which accordingly can be exhibited in the form

$$V = (z_1, z_2, \cdots),$$

where z_i is an ordinary 1-dimensional rational δ_i-cycle with $\delta_i \to 0$ and for any $e > 0$ there is an integer I such that for $i, j > I$, $z_i \underset{e}{\approx} z_j$.[9] Let $I = I_1$ be determined for the number $e = \epsilon$.

By uniform continuity of T there exists a $d > 0$ such that any set in A of diameter $< d$ transforms into a set of diameter $< \epsilon$ under T. Now by virtue of a theorem of Alexandroff's[8] there exists a $\delta > 0$ such that for any rational δ-cycle z in N there exists a rational Vietoris cycle W in N such that $W \underset{d}{\sim} z$.

Now since $\delta_i \to 0$ it follows from the theorem in §4 that there exists an integer I_2 such that if $i > I_2$ there exists a 1-dimensional rational δ-cycle C_i in N such that $f(C_i) \underset{\epsilon}{\sim} z_i$. Let us choose a fixed $k > I_1 + I_2$ and find C_k in N so that $f(C_k) \underset{\epsilon}{\sim} z_k$ and W in N so that $W \underset{d}{\sim} C_k$. We then have

$$f(W) \underset{\epsilon}{\sim} f(C_k) \underset{\epsilon}{\sim} z_k \underset{\epsilon}{\approx} z_{k+j} \qquad (j = 1, 2, \cdots).$$

Whence

$$f(W) \underset{\epsilon}{\sim} (z_k, z_{k+1}, \cdots) \sim V,$$

as required by our theorem.

6. **Mapping of Betti groups.** For any closed set F we let $p^r(F)$ denote the r-dimensional Betti number of F and let $B_R^r(F)$ denote the r-dimensional rational Betti group of F.[10] We begin this section with

[7] See Vietoris, Mathematische Annalen, vol. 97 (1927), pp. 454–572. For the rational Vietoris cycle (which we use) see Lefschetz, Annals of Mathematics, (2), vol. 29 (1928), pp. 232–254.

[8] See Alexandroff, Mathematische Annalen, vol. 106 (1932), pp. 161–238.

[9] This notation means that there exist rational numbers c_1 and c_2 such that $c_1 z_i - c_2 z_j$ bounds a rational e-complex in M. See Lefschetz, loc. cit., and Alexandroff, loc. cit.

[10] That is, $p^r(F)$ is the maximum number of r-dimensional rational Vietoris cycles (or true cycles) in F which are linearly independent relative to homologies in F and $B_R^r(F)$ is the homology class group of such cycles. See Lefschetz, loc. cit. It is immaterial which type of infinite cycle is used since, at least when $p^r(F)$ is finite, whichever is used in the definition, the resulting Betti groups are isomorphic.

(6.1) Lemma. *If F is compact and $p^r(F)$ is finite, there exists an $\epsilon > 0$ such that $\alpha^r \underset{\epsilon}{\sim} \beta^r$ (in F) implies $\alpha^r \sim \beta^r$ (in F) for true cycles (hence also for Vietoris cycles) α^r and β^r in F.*

Proof. Let

$$Z_1, Z_2, \cdots, Z_p \qquad (p = p^r(F))$$

be linearly independent true r-dimensional cycles in F. By a lemma of Alexandroff's[11] there exists an $\epsilon > 0$ such that $Z_1, Z_2, \cdots, Z_p$ are ϵ-independent in F. Now suppose $\alpha^r \underset{\epsilon}{\sim} \beta^r$ (in F), where α^r and β^r are true cycles in F. We have

$$\alpha^r \sim \sum_1^p a_i Z_i, \qquad \beta^r \sim \sum_1^p b_i Z_i,$$

since $p = p^r(F)$. From this we have $\sum a_i Z_i \sim \alpha^r \underset{\epsilon}{\sim} \beta^r \sim \sum b_i Z_i$, or $\sum (a_i - b_i) Z_i \underset{\epsilon}{\sim} 0$, whence $a_i \equiv b_i$. Accordingly, $\alpha^r - \beta^r \sim 0$ or $\alpha^r \sim \beta^r$.

Now, resuming the notation used in §§4 and 5, we have

(6.2) Theorem. *If $p^1(M)$ is finite, $B^1_R(N)$ maps homomorphically onto $B^1_R(M)$ under T.*

For let h denote the homomorphism of $B^1_R(N)$ into a subgroup of $B^1_R(M)$ generated by T. We have to show that $h[B^1_R(N)] = B^1_R(M)$, i.e., $B^1_R(N)$ maps onto all of $B^1_R(M)$. To this end let ϵ be determined as in (6.1) for the set $F = M$. Let H be any homology class of $B^1_R(M)$ and let V be any cycle in H. By §5 there exists a cycle W in N such that $f(W) \underset{\epsilon}{\sim} V$. Accordingly, $f(W) \sim V$ by the choice of ϵ. Whence $f(W) \in H$ so that if L denotes the homology class in $B^1_R(N)$ containing W we have $h(L) = H$.

(6.3) *If $p^1(N)$ is finite, $p^1(M)$ is finite and $\leqq p^1(N)$.*

Suppose, on the contrary, that there exists in M a system of linearly independent one-dimensional rational true cycles

$$V_1, V_2, \cdots, V_{p+1}, \qquad p = p^1(N).$$

Now it follows[11] that for some $\epsilon > 0$ these cycles are ϵ-independent relative to homologies in M. By uniform continuity of T there is a $\delta > 0$ such that any set in N of diameter $< \delta$ maps into a set of diameter $< \epsilon$. Now, by §5, for each $i \leqq p + 1$ there exists a true cycle W_i in N such that $f(W_i) \underset{\epsilon}{\sim} V_i$. Since $p^1(N) = p$, there must exist a relation of the form

$$a_1 W_1 + a_2 W_2 + \cdots + a_{p+1} W_{p+1} \sim 0,$$

whence

$$a_1 W_1 + a_2 W_2 + \cdots + a_{p+1} W_{p+1} \underset{\delta}{\sim} 0.$$

This gives

$$0 \underset{\epsilon}{\sim} \sum_1^{p+1} a_i f(W_i) \underset{\epsilon}{\sim} \sum a_i V_i$$

contrary to our supposition.

[11] Fundamenta Mathematicae, vol. 22 (1934), p. 18.

Clearly (6.2) and (6.3) yield at once

(6.4) $p^1(M) \leqq p^1(N)$; *and if either of these numbers is finite,* $B^1_R(N)$ *maps homomorphically onto* $B^1_R(M)$. · *In particular,*[12] *taking* $N = A$, *we have* $p^1(B) \leqq p^1(A)$.

7. **The cyclic case.** In case B is a cyclic continuum it follows by (3.2) that $T(N)$ must be equal to B for any non-degenerate nodal subset N of A. Hence the preceding results yield the following

THEOREM. *If A is a compact continuum, $T(A) = B$ is interior and B is cyclic, then* $p^1(B) \leqq p^1(N)$ *for any non-degenerate nodal subset N of A. Furthermore, if* $p^1(B)$ *or* $p^1(N)$ *(for any N) is finite,* $B^1_R(N)$ *maps homomorphically onto* $B^1_R(B)$.

In conclusion we remark that if A is a locally connected continuum all of the preceding results hold in particular if we choose for N any non-degenerate *node* of A, i.e., any true cyclic element of A containing just one cut point of A.

UNIVERSITY OF VIRGINIA.

[12] Compare with a result of Eilenberg's in Fundamenta Mathematicae, vol. 27 (1936), p. 163.

CERTAIN INTEGRAL FUNCTIONS RELATED TO EXPONENTIAL SUMS

By Harry Matison

Introduction. The study of Borel transforms by several mathematicians, in particular by Pólya,[1] has brought out many interesting relations between the singularities of a function $f(z)$ defined by a power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

on the one hand, and the rate of growth and distribution of zeros of its Borel transform

$$F(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n \tag{1}$$

on the other. As the prototype in these considerations, there is the case in which $f(z)$ is a rational function with simple poles

$$f(z) = \frac{\alpha_1}{1 - \beta_1 z} + \cdots + \frac{\alpha_k}{1 - \beta_k z},$$

and the Borel transform is a sum of exponentials

$$F(z) = \alpha_1 e^{\beta_1 z} + \cdots + \alpha_k e^{\beta_k z}.$$

This case has been studied in detail by Pólya and Schwengeler.[2] It is easily seen that here the coefficients a_n can be interpolated by the function

$$p(t) = \alpha_1 \beta_1^t + \cdots + \alpha_k \beta_k^t$$

with $a_n = p(n)$. In the particular case in which the β_ν all lie on the unit circle, $p(t)$ becomes a special type of almost-periodic function. The generalization to general uniformly almost-periodic functions (u. a. p. functions) has been studied by Bochner and Bohnenblust[3] as regards $f(z)$. We propose to make a

Received June 14, 1937; in revised form, January 6, 1938. This is essentially a dissertation accepted by the faculty of Princeton University for the degree of Doctor of Philosophy.

[1] G. Pólya, *Untersuchungen über Lücken und Singularitäten von Potenzreihen*, Math. Zeitschrift, vol. 29 (1929), p. 549.

[2] Pólya, *Geometrisches über die Verteilung der Nullstellen gewisser ganzer transzendenter Funktionen*, Münchner Berichte, 1920.

E. Schwengeler, Dissertation, Zürich, 1923.

[3] S. Bochner and F. Bohnenblust, *Analytic functions with almost periodic coefficients*, Annals of Math., vol. 35 (1934), p. 152.

corresponding study of the Borel transform $F(z)$. This we do in Part I. To be more precise, in Part I we discuss properties of the functions $F(z)$, where the a_n form an almost-periodic sequence. The methods used are found in part to be general enough to include the wider class

$$(2) \qquad F_\alpha(z) = \sum_{n=0}^{\infty} \frac{a_n}{\Gamma(\alpha n + 1)} z^n,$$

where, for reasons to become evident later, α is restricted to the range $0 < \alpha \leqq 2$.

In Part II we generalize the results of Pólya and Schwengeler in a different direction, namely, we study the functions

$$(3) \qquad G(z) = \sum_{n=1}^{\infty} B_n e^{\alpha_n z},$$

where the B_n form an absolutely convergent series and the α_n form a bounded set. Such functions have also been studied by M. Regensburger in a recent paper.[4] Despite certain similarities, our results are different from those of Regensburger, who follows closely the methods of Pólya and Schwengeler. Also in this connection we discuss an example of a function $G(z)$ which, unlike the rest of our results, has properties not to be expected from analogy with finite exponential sums.

The subject of this study was suggested by Professors S. Bochner and F. Bohnenblust. The writer takes pleasure in acknowledging his indebtedness for their many helpful discussions during its progress.

I

1. **Almost-periodic sequences.** Consider the additive abelian group of the rational integers and let $g(n)$ denote an almost-periodic function on this group.[5] The sequence $g(\nu)$ ($\nu = 0, 1, \cdots$) is then called an a. p. sequence. A. Walther has studied such sequences in detail, and we shall state those of his results which we shall need.[6]

We have the existence of the mean value

$$M\{g(n)\} = \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=0}^{n-1} g(\nu)$$

and the Fourier development

$$g(n) \sim \sum_{\nu=1}^{\infty} A_\nu e^{-i\lambda_\nu n},$$

where

$$A_\nu = M\{g(k)e^{i\lambda_\nu k}\}$$

[4] M. Regensburger, Math. Annalen, vol. 111 (1935), p. 505.

[5] J. von Neumann, Trans. Am. Math. Soc., vol. 36 (1934), p. 445.

[6] A. Walther, Atti Congresso Int., Bologna, 1928, vol. II, pp. 289–298.

are the Fourier coefficients and $-\lambda_\nu$ the corresponding exponents which are chosen so that $0 \leqq \lambda_\nu < 2\pi$.

With the given a.p. sequence $g(n)$ we can define a u.a.p. function $g(t)$ by linear interpolation as follows:[7]

$$g(t) = g(\nu) + \theta\{g(\nu + 1) - g(\nu)\}, \qquad \nu = [t], \qquad \theta = t - [t].$$

Conversely, the values taken by a u.a.p. function for integral values of the argument form an a.p. sequence. For the function just defined we have a Fourier series

$$g(t) \sim \sum_{k=1}^{\infty} c_k e^{-i\Lambda_k t}, \tag{4}$$

where the following relations hold:

$$\begin{aligned} A_\nu &= 0 \text{ if } \lambda_\nu \not\equiv \Lambda_k \pmod{2\pi} \qquad (k = 1, 2, \cdots),\\ A_\nu &= g_{\lambda_\nu}(0) \qquad \text{otherwise,} \end{aligned}$$

where $g_\varphi(t)$ is the periodic function whose Fourier series consists of all terms of (4) for which $\Lambda_k \equiv \varphi \pmod{2\pi}$. Thus we see that the $g(n)$ can be uniformly approximated by exponential sums $\sum_1^N d_\nu e^{-i\lambda_\nu n}$.

2. **Approximation to $F(z)$ by exponential sums.** If we approximate the sequence $g(n)$ uniformly within ϵ, we have

$$g(n) = \sum_1^N d_\nu^{(N)} e^{-i\lambda_\nu n} + \eta_n,$$

where $|\eta_n| < \epsilon$. Hence

$$F(z) = \sum_1^N d_\nu^{(N)} \exp(ze^{-i\lambda_\nu}) + \delta(z),$$

where $|\delta(z)| < \epsilon |e^z|$. Thus if $|z|$ is bounded, $F(z)$ is the uniform limit of a sequence of exponential sums

$$S_N(z) = \sum_{\nu=1}^N d_\nu^{(N)} \exp(ze^{-i\lambda_\nu}). \tag{5}$$

Similarly, $F_\alpha(z)$ is in any bounded region the uniform limit of a sequence

$$\mathfrak{S}_N(z) = \sum_{\nu=1}^N d_\nu^{(N)} E_\alpha(ze^{-i\lambda_\nu}), \tag{6}$$

where

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)} \tag{7}$$

is the function of Mittag-Leffler.

[7] With but few minor changes we could avoid bringing in u.a.p. functions and deal with a.p. sequences only.

3. **Representation of $F(z)$ by means of a u.a.p. function.** By means of the theory of multipliers of a Fourier series,[8] it is possible to represent $F(z)$ as well as $F_\alpha(z)$ as values of certain u.a.p. functions. The function

$$\Gamma_z(\lambda) = \exp(ze^{-i\lambda}),$$

where z is restricted to a bounded region, can be developed into an ordinary Fourier series which, on account of the differentiability of the function, is absolutely convergent uniformly in z:

$$\Gamma_z(\lambda) = \sum_{\nu=-\infty}^{+\infty} b_\nu(z)e^{-i\nu\lambda}.$$

Consider next the function

$$\gamma_z(t) = \sum_{\nu=-\infty}^{+\infty} b_\nu(z)g(\nu + t),$$

where $g(t)$ is, as before, the u.a.p. function which interpolates the $g(n)$ linearly. Approximating to $g(t)$ by an exponential polynomial $\sum_{\nu=1}^{N} c_\nu^{(N)}e^{-i\Lambda_\nu t}$, we see that $\gamma_z(t)$ can be uniformly approximated by the functions

$$\gamma_z^{(N)}(t) = \sum_{\nu=1}^{N} c_\nu^{(N)}\Gamma_z(\lambda_\nu)e^{-i\Lambda_\nu t}$$

and hence $\gamma_z(t)$ is u.a.p. for any z. Now $\gamma_z(0)$ is capable of approximation by the sums (5) so that, since $F(z)$ also has this property, we conclude $F(z) = \gamma_z(0)$. On the other hand we observe that since $\Gamma_0(\lambda_\nu) = 1$, also $g(t) = \gamma_0(t)$. The above argument shows further that

$$\gamma_z(t) \sim \sum_{\nu=1}^{\infty} c_\nu\Gamma_z(\lambda_\nu)e^{-i\Lambda_\nu t}.$$

For the functions $F_\alpha(z)$ everything is similar except that we use the multiplier $\Gamma_z(\lambda) = E_\alpha(e^{-i\lambda}z)$ instead of the one used above.

4. **Invariance of the class $F(z)$ under certain transformations.** We consider next a rather general type of linear transformation under which the class $F(z)$ remains invariant. Let $\sigma(t)$ be a (complex-valued) function, continuous and bounded over $-\infty < t < \infty$, and let $\psi(t)$ be a function of bounded variation in the same interval. Consider the transformation

$$F^*(z) = \int_{-\infty}^{\infty} F(z + \sigma(t))\, d\psi(t). \tag{8}$$

We shall prove concerning this transformation the following theorem.

[8] S. Bochner, Math. Annalen, vol. 102 (1929), pp. 489–504.

THEOREM I. *If $F(z)$ is given by a series* (1), *then $F^*(z)$ is also given by such a series. Its corresponding u.a.p. function which interpolates the coefficients is*

$$g^*(t) \sim \sum_1^\infty c_\nu \Gamma(\Lambda_\nu) e^{-i\Lambda_\nu t},$$

where $\Gamma(\Lambda)$ is a multiplier defined below.

Proof. We again make use of the notion of a multiplier of a Fourier series. Let

$$g_N(t) = \sum_{\nu=1}^N c_\nu^{(N)} e^{-i\Lambda_\nu t}$$

be a sequence of exponential polynomials approaching $g(t)$ uniformly. From a preceding paragraph we know that, in any finite region of the z-plane, $F(z)$ is uniformly approached by the sequence

$$S_N(z) = \sum_{\nu=1}^N c_\nu^{(N)} \exp(ze^{-i\Lambda_\nu}).$$

From (8) and the assumptions concerning $\sigma(t)$ and $\psi(t)$, it follows that $F^*(z)$ is approached in the same way by the sequence

$$S_N^*(z) = \sum_{\nu=1}^N c_\nu^{(N)} \Gamma(\Lambda_\nu) \exp(ze^{-i\Lambda_\nu}), \tag{9}$$

where

$$\Gamma(\Lambda) = \int_{-\infty}^{+\infty} \exp(\sigma(t)e^{-i\Lambda})\, d\psi(t).$$

If we now prove that there exists a u.a.p. function $g^*(t)$ which can be uniformly approximated by the sequence

$$g_N^*(t) = \sum_{\nu=1}^N c_\nu^{(N)} \Gamma(\Lambda_\nu) e^{-i\Lambda_\nu t}, \tag{10}$$

then Theorem I will be proved; for, the function (1) formed with $g^*(n)$ in place of $g(n)$ will be the uniform limit of the sequence (9). Now the function $\Gamma(\Lambda)$ clearly has two continuous derivatives in the interval $0 \leqq \Lambda \leqq 2\pi$ and also $\Gamma(\Lambda + 2\pi) = \Gamma(\Lambda)$. Hence we have a Fourier expansion

$$\Gamma(\Lambda) = \sum_{-\infty}^{\infty} d_\nu e^{-i\nu\Lambda},$$

where $\sum_{-\infty}^{\infty} |d_\nu|$ converges. It is now easy to show that the function

$$g^*(t) = \sum_{-\infty}^{\infty} d_\nu g(\nu + t)$$

is the one required. For, approximating $g(t)$ by exponential polynomials $g_N(t)$, we see that $g^*(t)$ can be approximated by the exponential polynomials (10).

Thus in terms of the u.a.p. function $g(t)$ the transformation (8) is described by introducing the multiplier $\Gamma(\Lambda_\nu)$ into the Fourier series.

We note the following special cases of (8):

(a) $\sigma(t) \equiv 0$, $\psi(t)$ arbitrary, $F^*(z) = cF(z)$;

(b) $d\psi(t) = 0$ except at a finite number of points, $\sigma(t)$ arbitrary,

$$F^*(z) = b_1F(z + \sigma_1) + \cdots + b_kF(z + \sigma_k);$$

(c) $\sigma(t) = e^{it}$, $\psi(t)$ differentiable in $0 \leqq t \leqq 2\pi$ and constant elsewhere,

$$F^*(z) = \frac{1}{2\pi i}\oint F(z + u)G(u)\,du$$

taken around the unit circle.

If we make a change of variables and choose $G(u)$ suitably, we have a special case of (c)

$$F^*(z) = \frac{n!}{2\pi i}\oint \frac{F(u)}{(u - z)^{n+1}}\,du = F^{(n)}(z).$$

Therefore, if $F(z)$ is defined by a series of the type (1), the same is true of all its derivatives. This is a fact which is also easily seen directly.

The transformation from $F(z)$ to $F^{(n)}(z)$ can be expressed very simply in terms of the u.a.p. function $g(t)$. The required transformation is, in fact, $g^*(t) = g(t + n)$.

Other transformations on $g(t)$ which leave the classes $F(z)$, $F_\alpha(z)$ invariant are, of course,

(i) $$g^*(t) = g_1(t)g_2(t),$$

where $g_1(t)$, $g_2(t)$ are two given u.a.p. functions,

(ii) $$g^*(t) = g_1(t) + g_2(t),$$

(iii) $$g^*(t) = \underset{(x)}{M}\{g_1(x + t), g_2(x)\} \qquad \text{(convolution)}.$$

It may be remarked that, in view of the results of Bochner and Bohnenblust, (i) gives rise to an interesting special case of Hadamard's theorem on the multiplication of singularities. For, $\sum_{n=0}^{\infty} g^*(n)z^n$ is the Hadamard product of $\sum_{n=0}^{\infty} g_1(n)z^n$ and $\sum_{n=0}^{\infty} g_2(n)z^n$. Since the exponents of the product of two u.a.p. functions are found by adding those of the factors, it follows that the singularities of the Hadamard product are contained among the products of singularities of the factors, which is just the result given by Hadamard's theorem.

The transformation (iii) may be interpreted as follows. By introducing zeros as coefficients if necessary, we can assume that $g_1(t)$ and $g_2(t)$ have the same exponents so that

$$g_1(t) \sim \sum A_k e^{-i\Lambda_k t}, \qquad g_2(t) \sim \sum B_k e^{-i\Lambda_k t}.$$

Then

$$g^*(t) \sim \sum A_k B_k e^{-i\Lambda_k t}.$$

Hence if we write, using an obvious notation,

$$F_{g_1}(z) \sim \sum A_k \exp (ze^{-i\lambda_k}), \qquad F_{g_2}(z) \sim \sum B_k \exp (ze^{-i\lambda_k}),$$

we have

$$F_{g^*}(z) \sim \sum A_k B_k \exp (ze^{-i\lambda_k}).$$

5. $F(z)$ **as a function of exponential type.** We next begin a more detailed study of a particular $F(z)$ and $F_\alpha(z)$. As for $F(z)$, it is clear, since the $g(n)$ are bounded, that it is an integral function, and moreover, one of the exponential type. For if $|g(n)| < c$ we have

$$|F(z)| \leqq \sum_{n=0}^{\infty} \frac{|g(n)|}{n!} |z|^n < ce^{|z|}.$$

This shows that the type is at most unity. We shall see below that the type is exactly unity.

Similarly, $F_\alpha(z)$ is an integral function of order $1/\alpha$ as may be seen with the aid of Stirling's formula to estimate the coefficients, $g(n)/\Gamma(\alpha n + 1)$. In this case also the type will be found to be unity.

In applying to our case the theory of functions of exponential type as developed by Pólya, we shall have to make use of properties of the Borel transform

(11) $$f(z) = \sum_0^{\infty} g(n)z^n$$

of $F(z)$. The result obtained by Bochner and Bohnenblust in their study of this function follows.

The set of singularities of $f(z)$ is exactly the closure of the point set $\{e^{i\lambda_k}\}$ where $-\lambda_k$ are the Fourier exponents of the a.p. sequence $g(n)$. If $f(z)$ is continued analytically over any regular point of the unit circle, the continuation is always given by

$$-\sum_{n=1}^{\infty} g(-n)z^{-n}.$$

Introducing the function

$$\varphi(z) = \frac{1}{z} f\left(\frac{1}{z}\right) = \sum_{n=0}^{\infty} g(n)z^{-n-1},$$

we have the fundamental formulas[9]

(12) $$F(z) = \frac{1}{2\pi i} \oint \varphi(\zeta)e^{z\zeta}\, d\zeta.$$

(13) $$\varphi(z) = \int_0^{\infty} F(\zeta)e^{-z\zeta}\, d\zeta,$$

[9] G. Pólya, Math. Zeitschrift, vol. 29 (1929), p. 580.

the first integral being taken around a contour which contains the unit circle in its interior. The second integral is taken along the positive real axis, but in this form converges only in a half-plane. By rotating the path of integration through a suitable angle, the formula (13) can be used for any value of z. We have similar formulas for $F_\alpha(z)$; namely, if we write $\beta = \alpha^{-1}$,

$$\frac{1}{2\pi i}\oint \varphi(\zeta)E_\alpha(z\zeta)\,d\zeta = \frac{1}{2\pi i}\oint \sum_{n=0}^{\infty}\sum_{\nu=0}^{\infty}\frac{g(n)}{z^{n+1}}\frac{z^\nu\zeta^\nu\,d\zeta}{\Gamma(\alpha\nu+1)} = \sum_{k=0}^{\infty}\frac{g(k)z^k}{\Gamma(\alpha k+1)},$$

$$\int_0^\infty F_\alpha(\zeta)\exp\left(-(z\zeta)^\beta\right)\zeta^{\beta-1}\,d\zeta = \int_0^\infty \sum_{\nu=0}^{\infty}\frac{g(\nu)\zeta^\nu}{\Gamma(\alpha\nu+1)}\exp\left(-(z\zeta)^\beta\right)\zeta^{\beta-1}\,d\zeta$$

$$= \sum_{\nu=0}^{\infty}\frac{g(\nu)}{\Gamma(\alpha\nu+1)}\int_0^\infty \zeta^{\nu+\beta-1}\exp\left(-(z\zeta)^\beta\right)d\zeta = \alpha z^{1-\beta}\varphi(z).$$

Hence

$$\frac{1}{2\pi i}\oint \varphi(\zeta)E_\alpha(z\zeta)\,d\zeta = F_\alpha(z), \tag{14}$$

$$\int_0^\infty F_\alpha(\zeta)e^{-(z\zeta)^\beta}\zeta^{\beta-1}\,d\zeta = \alpha z^{1-\beta}\varphi(z). \tag{15}$$

The necessary interchanges of summation and integration are allowed because of the uniform and absolute convergence of the series. In (15) we must restrict z to the region of convergence of the integral, and, by rotating the line of integration in a manner similar to that used by Pólya for (13), we can obtain (15) for any other value of z. We shall, however, make no use of either (13) or (15) and so we omit these details.

6. **Radial growth of $F_\alpha(z)$.** We consider next the rate of growth of $F_\alpha(z)$ in a direction $z = re^{i\varphi}$ and define the function

$$A(\varphi) = A_k, \text{ if } \varphi = \lambda_k \text{ for some } k;$$
$$A(\varphi) = 0, \quad \text{if } \varphi \neq \lambda_k \quad \text{for all } k.$$

For fixed φ, $g(n) - A(\varphi)e^{-in\varphi}$ is an a.p. sequence since the a.p. property is invariant under addition. Let $\sigma_N(n) = \sum_{\nu=1}^{N}\delta_\nu^{(N)}e^{-i\mu_\nu n}$ be an ϵ-approximation to this sequence so that $0 \leqq \mu_\nu < 2\pi$, $\varphi \neq \mu_\nu$ for all ν and

$$|\, g(n) - A(\varphi)e^{-in\varphi} - \sigma_N(n)\,| < \epsilon.$$

Hence

$$F_\alpha(re^{i\varphi}) = A(\varphi)E_\alpha(r) + \sum_{\nu=1}^{N}\delta_\nu^{(N)}E_\alpha(e^{-i\mu_\nu}z) + G(z),$$

where $|\,G(z)\,| < \epsilon\, E_\alpha(r)$.

In order to proceed further we require the following facts concerning the growth of the Mittag-Leffler function $E_\alpha(z)$.[10]

$$(16)\qquad \begin{aligned} E_\alpha(z) &= \beta \exp(z^\beta) + O\left(\frac{1}{r}\right), \quad \text{if} \quad |\arg z| \leqq \frac{\alpha\pi}{2}; \qquad \beta = \frac{1}{\alpha}; \\ E_\alpha(z) &= -\frac{1}{z}\,\frac{1}{\Gamma(1-\alpha)} + O\left(\frac{1}{r^2}\right), \quad \text{if} \quad \pi \geqq |\arg z| > \frac{\alpha\pi}{2}. \end{aligned}$$

Returning to $F_\alpha(z)$, we have

$$\frac{F_\alpha(re^{i\varphi})}{E_\alpha(r)} = A(\varphi) + \sum_{\nu=1}^{N} \delta_\nu^{(N)} \frac{E_\alpha(re^{i(\varphi-\mu_\nu)})}{E_\alpha(r)} + \frac{G(z)}{E_\alpha(r)}.$$

However, since $\mu_\nu \not\equiv \varphi$, it is evident from (16) that, as $r \to \infty$, the second term on the right tends to zero, while from the approximation, the last term is in absolute value $< \epsilon$. Although the case $\alpha = 1$ does not fall under (16), the same statements nevertheless hold from well-known growth properties of $E_1(z) = e^z$. (A similar remark applies when $\alpha = z$.) Hence we have the result

$$(17)\qquad \lim_{r\to\infty} \frac{F_\alpha(re^{i\varphi})}{E_\alpha(r)} = A(\varphi) \qquad (0 < \alpha \leqq 2),$$

or, since $E_\alpha(r) \sim \beta e^{r^\beta}$,

$$(18)\qquad F_\alpha(re^{i\varphi}) \sim \beta A(\varphi) e^{r^\beta} \qquad (0 < \alpha \leqq 2).$$

It may be remarked that this gives a means of finding the Fourier series of the a.p. sequence $g(n)$ in terms of the growth properties of $F_\alpha(z)$.

Also this result allows us to conclude the above mentioned fact that $F_\alpha(z)$ is of order β and type unity.

7. **The Phragmén-Lindelöf function** $h(\varphi)$.[11] The Phragmén-Lindelöf function $h(\varphi)$, which we now introduce, is defined generally for a function $\Phi(z)$ which is regular and of order ρ in an angle $\varphi_1 < \varphi < \varphi_2$ as follows:

$$(19)\qquad h(\varphi) = \overline{\lim_{r\to\infty}} \frac{\log|\Phi(re^{i\varphi})|}{r^\rho} \qquad (\varphi_1 < \varphi < \varphi_2).$$

This function, called by Pólya the *indicator* of $\Phi(z)$, plays an important rôle in the following work. Therefore, we shall consider it in some detail. As we see from (19), $h(\varphi)$ measures the rate of growth of $\Phi(z)$ in the direction $re^{i\varphi}$.

For functions of exponential type, Pólya has given the following method for determining $h(\varphi)$. Let $\sum_{n=0}^{\infty} a_n z_n/(n!)$ be an integral function of exponential type and denote by $\mathfrak{J}$ the smallest closed convex region which contains all the singu-

[10] L. Bieberbach, *Lehrbuch der Funktionentheorie*, vol. II, p. 267.

[11] Acta Mathematica, vol. 31 (1908).

larities of the Borel transform $\sum_{n=0}^{\infty} a_n z^{-n-1}$. This region is called the *conjugate diagram* of the given integral function, and $\mathfrak{J}$, its reflection in the real axis, is called the *indicator diagram* of the given integral function. As Pólya shows, the supporting function (Stützfunktion) of $\mathfrak{J}$ is $h(\varphi)$.

Applying this to $F(z)$ we find, using the result of Bochner and Bohnenblust mentioned above, that $\mathfrak{J}$ is the smallest convex region containing the points $e^{i\lambda_k}$. Thus, if there is an arc of the unit circle on which the points $e^{i\lambda_k}$ lie everywhere dense, then this arc will form part of the boundary of $\mathfrak{J}$. Besides such arcs, $\mathfrak{J}$ will in general contain chords joining points on the unit circle.

We shall present a method of obtaining $h(\varphi)$ for $F(z)$ which, although similar to that of Pólya, makes use of the special nature of $F(z)$ and is, at least partially, applicable to $F_\alpha(z)$ for $\alpha \neq 1$. We consider first $F(z)$.

We start from the formula (12) and deform the path of integration around the singularities of $\varphi(\zeta)$ so that the new path is composed of contours running very near to the singularities and of radii of the unit circle each described twice in opposite directions. These and the small circle around the origin can be dropped since, by the result of Bochner and Bohnenblust, $\varphi(\zeta)$ is single-valued and bounded near $\zeta = 0$. We have from (12)

$$|F(Re^{i\varphi})| \leqq K \max |e^{z\zeta}| = K \max \{\exp[rR\cos(\theta + \varphi)]\},$$

where K depends upon the path of integration but is independent of R, and where the maximum is taken as $\zeta = re^{i\theta}$ runs over this path. Hence

$$h(\varphi) = \overline{\lim_{R\to\infty}} \frac{\log |F(Re^{i\varphi})|}{R} \leqq \max r\cos(\theta + \varphi).$$

But the path of integration runs very near to the singularities, S, so that we can pass now to the limit and have

$$h(\varphi) \leqq \max_S \{r\cos(\theta + \varphi)\} = \max_{\bar{S}} \{r\cos(\theta - \varphi)\},$$

the maximum being taken as ζ runs over the set $\bar{S}$ of conjugate points of the singularities of $\varphi(\zeta)$.

We next obtain the inequality in the other direction and for this purpose we need the following general property of $h(\varphi)$ due to Phragmén and Lindelöf. *If $h(\varphi)$ is the indicator of a function which is regular and of order ρ in a sector containing the angle $\varphi_1 < \varphi < \varphi_3$ and if $\varphi_1 < \varphi_2 < \varphi_3$, $\varphi_2 - \varphi_1 < \pi\rho^{-1}$, $\varphi_3 - \varphi_2 < \pi\rho^{-1}$, then*

$$h(\varphi_1)\sin\rho(\varphi_3 - \varphi_2) + h(\varphi_2)\sin\rho(\varphi_1 - \varphi_3) + h(\varphi_3)\sin\rho(\varphi_2 - \varphi_1) \geqq 0. \tag{20}$$

For $\rho = 1$ this means, essentially, that $h(\varphi)$ is the supporting function of a convex region. For $\rho \neq 1$ there is no such simple geometrical interpretation.

Coming back to $F(z)$, let $e^{i\psi}$ be a point of $\bar{S}$ which gives the largest projection on the given ray $Re^{i\varphi}$, i.e., for which the maximum above considered is attained,

and suppose that $h(\varphi) < \cos(\psi - \varphi) - \delta, \delta > 0$. We shall show that this supposition leads to a contradiction. Without sacrificing generality we may suppose $\psi < \varphi$. Take a third angle θ such that $\theta < \psi$, $(\theta - \varphi) < \pi$. Applying (20) we get, since $h(\psi) = 1$,

$$h(\theta)\sin(\varphi - \psi) + h(\varphi)\sin(\psi - \theta) + h(\psi)\sin(\theta - \varphi) \geqq 0,$$

$$h(\theta) > \cos(\psi - \theta) + \frac{\delta}{\sin(\varphi - \psi)}\sin(\psi - \theta).$$

This is impossible since the right side assumes values which are greater than unity for φ, ψ fixed and variable θ sufficiently near ψ.

Combining this result with the previously obtained inequality, we have

$$h(\varphi) = \max_{\bar{S}}\{\cos(\theta - \varphi)\}. \tag{21}$$

Geometrically this equation may be interpreted as follows. Consider the totality of circles with diameters OP, where O is the origin and P runs over $\bar{S}$. Given any direction φ, $h(\varphi)$ is measured by the largest intercept cut off on the ray $Re^{i\varphi}$ by these circles. It is easily seen that this is merely another way of saying that $h(\varphi)$ is the supporting function of the indicator diagram of $F(z)$.

To carry out a similar discussion for $F_\alpha(z)$, we start from the integral formula (14) and find

$$|F_\alpha(Re^{i\varphi})| \leqq K \max|E_\alpha(rRe^{i(\theta+\varphi)})|$$

as $\zeta = re^{i\theta}$ runs over the path of integration. Hence

$$h(\varphi) \leqq \overline{\lim_{R\to\infty}} \frac{\log\{\max|E_\alpha(rRe^{i(\theta+\varphi)})|\}}{R^\beta}.$$

But, from the growth properties (16) of $E_\alpha(z)$, we have

$$\begin{aligned} \overline{\lim_{R\to\infty}} \frac{\log|E_\alpha(Rre^{i(\theta+\varphi)})|}{R^\beta} &= r^\beta \cos\beta(\theta + \varphi) \qquad (|\theta + \varphi| \leqq \tfrac{1}{2}\alpha\pi), \\ &= 0 \qquad (\pi \geqq |\theta + \varphi| > \tfrac{1}{2}\alpha\pi). \end{aligned}$$

Hence, using, as before, the conjugate set $\bar{S}$, we find

$$h(\varphi) \leqq \max\{\max_{\bar{S}}[r^\beta \cos\beta(\theta + \varphi)], 0\}.$$

For those angles φ for which $h(\varphi) \geqq 0$, we can proceed exactly as before to obtain the inequality in the other direction so that in this case

$$h(\varphi) = \max_{\bar{S}}[\cos\beta(\varphi - \theta)].$$

The geometrical interpretation is analogous to that for $\alpha = 1$ except that here the curves $h = \cos\beta(\varphi - \theta)$ instead of circles $h = \cos(\varphi - \theta)$ are used. For values of φ for which $h(\varphi) < 0$ the method fails. That $h(\varphi)$ can be negative for functions of our class is shown by the example $\frac{1}{2}E_{\frac{1}{2}}(z) + \frac{1}{2}E_{\frac{1}{2}}(-z) = e^{z^2}$. Here

$h(\varphi) = \cos 2\varphi$. However, if we make the restriction that there be no arc of the unit circle greater than $\alpha\pi$ which is free from points $e^{i\lambda_k}$, it can easily be concluded that $h(\varphi)$ is never negative, since in this case $\max_{\bar{s}} \cos \beta(\varphi - \theta) \geqq 0$.

8. A theorem on the zeros of $F_\alpha(z)$.

THEOREM II. *If $F_\alpha(z)$ is not of the form ke^{cz^n}, where k and c are constants, $|c| = 1$, and n is a positive integer, then the series*

$$\sum_{\nu=1}^{\infty} r_\nu^{-\beta}, \tag{22}$$

where $r_\nu e^{i\theta_\nu}$ ($\nu = 1, 2, \cdots$) are the zeros of $F_\alpha(z)$, is divergent.

Remark. We know from the theory of integral functions that $\sum_{\nu=1}^{\infty} r_\nu^{-\rho}$ converges for every $\rho > \beta$. Our theorem shows that for our functions, with the exceptions mentioned, this is no longer true for $\rho = \beta$.

Proof. Suppose the series (22) were convergent. Then, by a well-known theorem,[12] we would have

$$F_\alpha(z) = e^{g(z)} z^q P(z),$$

where $P(z)$ is a Weierstrassian canonical product which belongs to the *minimal* type of order β and $g(z)$ is a polynomial of degree $[\beta]$ at most. If β were not integral, we could conclude that $F_\alpha(z)$ is of minimal type, contrary to the fact proved above that it is of type unity. Hence, suppose $\beta = [\beta] = n$. Then $g(z)$ is necessarily of degree n and

$$\frac{\log |F_\alpha(Re^{i\varphi})|}{R^n} = \frac{\Re\{g(Re^{i\varphi})\}}{R^n} + \frac{q \log R}{R^n} + \frac{\log |P(Re^{i\varphi})|}{R^n}.$$

If we take the limit superior, the last two terms give no contribution and we conclude that $h(\varphi) = \cos n(\varphi - \varphi_0)$ for a suitable φ_0 which, by a rotation of the z-plane, we may take equal to zero. We now have to determine which functions $F_\alpha(z)$ have

$$h(\varphi) = \cos n\varphi.$$

In the first place, there are exactly n directions of maximal growth; namely, $\varphi = 2k\pi/n$ ($k = 0, 1, \cdots, n-1$) and hence by (17), the a.p. sequence of the coefficients has exactly n Fourier exponents which are $0, -2\pi/n, \cdots, -2(n-1)\pi/n$. Our function is therefore of the form

$$F_\alpha(z) = A_0 E_\alpha(z) + A_1 E_\alpha(\eta_1 z) + \cdots + A_{n-1} E_\alpha(\eta_{n-1} z),$$

where $\eta_k = -2k\pi i/n$. Now, since each of these summands cannot grow exponentially small, there must exist certain conditions on the A_k which enable

[12] L. Bieberbach, *Funktionentheorie*, vol. II, p. 242.

the sum to do so. In order to find these conditions, we make use of the following result on the growth of $E_\alpha(z)$.[13]

$$E_{1/n}(z) = \sum_{\nu=1}^{n-1} \frac{z^{-\nu}}{\Gamma\left(-\frac{\nu}{n}+1\right)} + O(|z|^{-n}), \qquad |\arg z| > \frac{\pi}{2n}.$$

The required condition which the A_ν must satisfy is that the coefficients of the various negative powers of z must vanish. This gives the equations

$$A_0 + A_1\epsilon_1 + A_2\epsilon_2 + \cdots + A_{n-1}\epsilon_{n-1} = 0,$$
$$A_0 + A_1\epsilon_1^2 + A_2\epsilon_2^2 + \cdots + A_{n-1}\epsilon_{n-1}^2 = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$A_0 + A_1\epsilon_1^{n-1} + A_2\epsilon_2^{n-1} + \cdots + A_{n-1}\epsilon_{n-1}^{n-1} = 0,$$

where $1, \epsilon_1, \cdots, \epsilon_{n-1}$ are the distinct n-th roots of unity. Taking A_0 as given, we see that $A_1, \cdots, A_{n-1}$ are uniquely determined since the determinant of these equations is a Vandermonde determinant and is thus different from zero. But since $A_0 = A_1 = \cdots = A_{n-1}$ is a solution of the system, it is the only one. Hence

$$F_\alpha(z) = A_0[E_{1/n}(z) + \cdots + E_{1/n}(\epsilon_{n-1}z)] = A_0e^{z^n}.$$

This completes the proof.

9. **The zeros of $F_\alpha(z)$ in certain sectors.** We next consider the distribution of the zeros of $F_\alpha(z)$ in a sector S, $\varphi_1 < \arg z < \varphi_2$, such that the interval $\varphi_1 < \varphi < \varphi_2$ is covered everywhere densely by the points λ_ν.

By the result (18) on the growth of $F_\alpha(z)$, together with the fact that $h(\varphi)$ is continuous,[14] we have $h(\varphi) = 1$, $\varphi_1 \leqq \varphi \leqq \varphi_2$, and also

$$\lim_{R\to\infty} \frac{\log|F_\alpha(Re^{i\varphi})|}{R^\beta} = 1 \tag{23}$$

for a set of φ everywhere dense in this interval.

Suppose $F_\alpha(z) \neq 0$ in the sector S. Then

$$G_\alpha(z) = [F_\alpha(z)]^{-1}$$

would be regular in S. Let S' be a slightly smaller sector. Then a theorem on the minimum modulus of an integral function[15] tells us that $G_\alpha(z)$ is of order β in S'. Let $H(\varphi)$ denote the indicator of $G_\alpha(z)$ in S'. By (23) and by the continuity of $H(\varphi)$, we have $H(\varphi) = -1$ in S'. Now take three directions

[13] A. Wiman, Acta Mathematica, vol. 29 (1905), p. 199.
[14] The continuity of $h(\varphi)$ follows easily from (20).
[15] See Bieberbach, *Funktionentheorie*, vol. II, p. 261.

θ, $\theta + \eta$, $\theta + 2\eta$, where $\eta > 0$, all lying in this sector. Applying the inequality (20), we get

$$(-1)\sin\beta\eta - (-1)\sin 2\beta\eta + (-1)\sin\beta\eta \geqq 0,$$

$$2\sin\beta\eta \leqq \sin 2\beta\eta.$$

This is false for η sufficiently small. Hence $F_\alpha(z)$ must have at least one zero in S. It follows easily that $F_\alpha(z)$ has an infinite number of zeros in S; for, if there were only a finite number, we could apply the foregoing argument to the function

$$Q_\alpha(z) = F_\alpha(z)(z - z_1)^{-1}(z - z_2)^{-1} \cdots (z - z_k)^{-1},$$

where $z_1, \cdots, z_k$ are all the zeros of $F_\alpha(z)$ in S, and arrive at the contradictory conclusion that $Q_\alpha(z)$ has at least one zero in S.

By a refinement of the above method we can go much farther; namely, we can show that the series

$$\sum_{\nu=1}^{\infty} |z_\nu|^{-\beta} \tag{24}$$

over the zeros in S is divergent. However, we first require two lemmas.

LEMMA 1. *If* $\sum_{\nu=1}^{\infty} |a_\nu|^{-\rho}$ *converges and if* $P(z)$ *is the canonical product formed with the zeros* a_ν, *then for any* $\epsilon > 0$

$$|P(z)| > e^{-\epsilon|z|^\rho}$$

for sufficiently large $|z|$ *satisfying the condition* $|1 - z/a_\nu| > \delta > 0$.

Proof. The proof is similar to that which gives the well-known inequality in the other direction.[16] We have by definition

$$P(z) = \prod_{n=1}^{\infty}(1 - \zeta_n)\exp\left(\zeta_n + \frac{1}{2}\zeta_n^2 + \cdots + \frac{1}{q}\zeta_n^q\right) = \prod_1^\infty \mathfrak{E}(\zeta_n, q), \qquad \zeta_n = \frac{z}{a_n},$$

where q is the integer such that $q < \rho \leqq q + 1$. Now if $|\zeta_n| < \frac{1}{2}$,

$$|\mathfrak{E}(\zeta_n, q)| = \left|\exp\left\{-\frac{1}{q+1}\zeta_n^{q+1} - \cdots\right\}\right| > \exp\left\{\frac{|\zeta_n^{q+1}|}{|\zeta_n| - 1}\right\} > e^{-2\zeta_n^\rho}.$$

For $|\zeta_n|$ large, say $> \lambda$, the same inequality holds since $\rho > q$. For $\frac{1}{2} \leqq |\zeta_n| \leqq \lambda$, we can find a positive number A such that

$$|\mathfrak{E}(\zeta_n, q)| > e^{-A|\zeta_n|^\rho}. \tag{25}$$

Also we can take $A > 2$ so that we have (25) holding for all $|\zeta_n|$. Now choose N so that

$$\sum_{n=N+1}^{\infty} |a_n|^{-\rho} < \frac{\epsilon}{2A}$$

[16] Bieberbach, *Funktionentheorie*, vol. II, p. 239.

and write

$$P(z) = \prod_1^N \cdot \prod_{N+1}^{\infty} \mathfrak{E}(\zeta_n, q).$$

Given $\epsilon > 0$, we have for sufficiently large $|z|$

$$\left| \prod_1^N \right| > \exp\left\{ -\frac{\epsilon}{2} |z|^{\rho} \right\}.$$

Also, by (25),

$$\left| \prod_{N+1}^{\infty} \right| > \exp\left\{ -A|z|^{\rho} \sum_{N+1}^{\infty} |a_n|^{-\rho} \right\} > \exp\left(-\frac{\epsilon}{2} |z|^{\rho} \right).$$

This proves the lemma.

LEMMA 2. *If the a_ν of Lemma 1 all lie in a sector S, then for any direction $re^{i\varphi}$ outside S*

$$\lim_{r \to \infty} \frac{\log |P(re^{i\varphi})|}{r^{\rho}} = 0.$$

Proof. A well-known theorem on canonical products tells us that

$$\overline{\lim_{r \to \infty}} \frac{\log |P(re^{i\varphi})|}{r^{\rho}} \leqq 0.$$

Now along any ray outside S it is clear that $|1 - z/a_n| > \delta$ for some $\delta > 0$ and all n. Hence, Lemma 1 gives

$$\underline{\lim_{r \to \infty}} \frac{\log |P(re^{i\varphi})|}{r^{\rho}} \geqq 0.$$

We can now return to $F_\alpha(z)$. Assume that the series (24) taken over the zeros of $F_\alpha(z)$ in S converges. Write

$$F_\alpha(z) = P(z)Q(z),$$

where $P(z)$ is the canonical product formed with the zeros z_ν. Let ψ denote a direction in S along which (23) holds. We wish to show that a similar relation holds for $Q(z)$ along the same ray. We have first

$$\underline{\lim_{r \to \infty}} \frac{\log |Q(Re^{i\psi})|}{R^{\beta}} \geqq 1.$$

Otherwise since

$$\overline{\lim_{r \to \infty}} \frac{\log |P(Re^{i\psi})|}{R^{\beta}} \leqq 0,$$

we should have a sequence of points along the ray contradicting (23).

For the second part of the proof, we require the inequality of Phragmén and Lindelöf (20), but in a slightly different form. *Let $\alpha < \theta_1 < \theta_2 < \beta$ and $\theta_2 -$*

$\theta_1 < \pi$, and let $h(\theta_1) \leqq h_1$, $h(\theta_2) \leqq h_2$. Let $H(\theta)$ be the function of the form $a \cos \theta + b \sin \theta$ which takes the values h_1, h_2 at θ_1, θ_2, respectively. Then

$$h(\theta) \leqq H(\theta) \qquad (\theta_1 \leqq \theta \leqq \theta_2).$$

Suppose now that for some $\delta > 0$

$$\overline{\lim_{R\to\infty}} \frac{\log | Q(Re^{i\psi}) |}{R^\beta} = \lambda > 1 + \delta.$$

Choose two rays σ, τ on either side of ψ and near enough to it so that the $H(\theta)$ which is equal to 1 for σ, τ is necessarily $< 1 + \delta$ in between. Choose two more rays ξ, η so that $\sigma < \xi < \psi < \eta < \tau$. Now write

$$F_\alpha(z) = P(z)Q(z) = P_1(z)P_2(z)Q(z) = P_1(z)Q_1(z),$$
$$Q_1(z) = P_2(z)Q(z),$$

where $P_1(z)$ is formed from the zeros in the sector ξ, η and $P_2(z)$ from the remaining zeros. By Lemma 2 we have along both the rays σ and τ

$$\overline{\lim_{R\to\infty}} \frac{\log | Q_1(z) |}{R^\beta} \leqq \overline{\lim_{R\to\infty}} \frac{\log | F_\alpha(z) |}{R^\beta} + \lim_{R\to\infty} \frac{\log | P_1(z) |}{R^\beta} = 1.$$

Hence, by the choice of σ, τ and by the inequality of Phragmén-Lindelöf,

$$\overline{\lim_{R\to\infty}} \frac{\log | Q_1(Re^{i\psi}) |}{R^\beta} < 1 + \delta.$$

Again by Lemma 2, since $P_2(z)$ has no zeros in the sector ξ, η, we have

$$\overline{\lim_{R\to\infty}} \frac{\log | Q(Re^{i\psi}) |}{R^\beta} \leqq \overline{\lim_{R\to\infty}} \frac{\log | Q_1(Re^{i\psi}) |}{R^\beta} + \lim_{R\to\infty} \frac{\log | P_2(Re^{i\psi}) |}{R^\beta} < 1 + \delta.$$

This contradicts the above assumption. Hence we have arrived at the result similar to (23),

$$\lim_{R\to\infty} \frac{\log | Q(Re^{i\psi}) |}{R^\beta} = 1.$$

But (23) was all that was required in order to prove that $F_\alpha(z)$ has at least one zero in S. Hence, we can proceed in the same way to the conclusion that $Q(z)$ has at least one zero in S. This contradicts the assumption that all the zeros of $F_\alpha(z)$ were included in $P(z)$. This completes the proof of the following theorem.

THEOREM III. *If S is a sector which cuts off from the unit circle an arc on which the points $e^{i\lambda_k}$ lie everywhere densely, then the series*

$$\sum_{\nu=1}^{\infty} | z_\nu |^{-\beta},$$

taken over the zeros of $F_\alpha(z)$ in S, is divergent.

10. The zeros along the normal to a side of $\mathfrak{Z}$. We consider in this paragraph the case in which the indicator diagram contains a chord joining two isolated points of the set $\{e^{i\lambda_k}\}$, these considerations being, however, confined to the case $\alpha = 1$. Without loss of generality we can assume that these points are $e^{i\mu}$ and $e^{-i\mu}$, so that we have symmetry about the x-axis. In this case it is not difficult, following the method of Pólya for finite exponential sums, to discuss the distribution of the zeros of $F(z)$ in the sector S, $-\mu \leqq \arg z \leqq \mu$.

We can write

$$F(z) = A \exp(ze^{i\mu}) + B \exp(ze^{-i\mu}) + \Psi(z),$$

where A, B are the Fourier coefficients of $g(n)$ corresponding to $-\mu$, μ, respectively. By comparing the indicator diagram of $\Psi(z)$ with that of $F(z)$, we easily see that in the sector $\eta \leqq \arg z \leqq \mu$, where η is small and positive, the growth of $F(z)$ is dominated by that of $\exp(ze^{-i\mu})$. Similarly, in the sector $-\eta \geqq \arg z \geqq -\mu$, $F(z)$ is dominated by $\exp(ze^{i\mu})$. Hence, of the zeros of $F(z)$ in S, almost all are to be found in any sector S_η, $-\eta < \arg z < \eta$.

In order to study the zeros in S_η, we make use of the theorem of Rouché. Along the sides of S_η, $F(z)$ is dominated by the function

$$\Phi(z) = A \exp(ze^{i\mu}) + B \exp(ze^{-i\mu}),$$

as we have already seen. We have

$$\Phi(z) = B \exp(ze^{-i\mu}) \left\{\frac{A}{B} + e^{i\sigma z}\right\} \qquad (\sigma = 2 \sin \mu),$$

so that the zeros of $\Phi(z)$ are given by

$$z = \frac{1}{\sigma} \log \frac{B}{A} + \frac{2k\pi}{\sigma} \qquad (k = 0, \pm 1, \pm 2, \cdots),$$

and are therefore spaced at equal intervals along a line parallel to the real axis. Also, if we cut out from S_η small circles of a fixed radius around each of these zeros of $\Phi(z)$, in the remainder, for suitable K independent of z, we have

$$|\Phi(z)| > K |\exp(ze^{-i\mu})|, \qquad 0 \leqq \arg z \leqq \eta,$$

$$|\Phi(z)| > K |\exp(ze^{i\mu})|, \qquad 0 \geqq \arg z \geqq \eta.$$

Hence we can find a sequence of equidistant radii $R_\nu \to \infty$ such that for $z_\nu = R_\nu e^{i\varphi}$, $-\eta \leqq \varphi \leqq \eta$, we have

$$\lim_{\nu \to \infty} \frac{F(z_\nu) - \Phi(z_\nu)}{\Phi(z_\nu)} = 0.$$

Along the sides of S_η, we already have $F(z) \sim \Phi(z)$. Hence, outside a sufficiently large circle, the number of zeros of $F(z)$ in S_η is the same as the number of zeros of $\Phi(z)$. If we use $N(r)$, $N_1(r)$ to denote respectively the enumeration functions of the zeros of $F(z)$ and $\Phi(z)$, we therefore have the result that the

difference $N_1(r) - N(r)$ remains bounded. Moreover, since $N_1(r) = \sigma r/(2\pi) + O(1)$, we have

$$N(r) = \frac{\sigma r}{2\pi} + O(1).$$

We note that σ is the length of the side of the indicator diagram joining $e^{i\mu}$ and $e^{-i\mu}$.

It may be remarked that the preceding result holds also if the points $e^{\pm i\mu}$ are endpoints of arcs of the indicator diagram, provided we assume that the terms of the Fourier series corresponding to all points in some neighborhood of each of $e^{\pm i\mu}$ points form an absolutely convergent series. For, it is easily seen that the growth of $F(z)$ in the sector S depends only upon the terms corresponding to points in any neighborhood of the endpoints $e^{\pm i\mu}$ and these terms are in turn dominated by $\Phi(z)$ as can be shown under the hypothesis of absolute convergence. However, since the absolutely convergent case has been considered by Regensburger, we shall not go into details here. Some aspects of the absolutely convergent case are also considered in Part II.

II

In this part we consider functions defined by series

$$G(z) = \sum_{k=1}^{\infty} B_k e^{\alpha_k z}, \tag{1}$$

where $\sum_{1}^{\infty} | B_k |$ converges and $| \alpha_k | \leqq A$, $A > 0$ $(k = 1, 2, \cdots)$. This class of functions contains those $F(z)$ considered in Part I for which the u.a.p. function $g(t)$ has an absolutely convergent Fourier series.

It is trivial that $G(z)$ is an integral function of exponential type since

$$| G(z) | \leqq \left(\sum_{k=1}^{\infty} | B_k | \right) \cdot e^{A|z|}.$$

We therefore introduce the Borel transform, $\gamma(z)$, of $G(z)$. We have

$$\gamma(z) = \int_0^{\infty} G(\zeta) e^{-z\zeta} \, d\zeta = \sum_{\nu=1}^{\infty} B_\nu \int_0^{\infty} e^{(\alpha_\nu - z)\zeta} \, d\zeta = \sum_{\nu=1}^{\infty} \frac{B_\nu}{z - \alpha_\nu}. \tag{2}$$

In order to study the growth of $G(z)$ in different directions by means of its indicator diagram, we must first learn something about the situation of the singularities of $\gamma(z)$. From (2) one would expect that these would consist of the set $\{\alpha_k\}$ and its limit points. In general, however, this is not the case as has been shown by examples due to Wolff and Pringsheim.[17] It is clear that $\gamma(z)$ is regular outside the closure of the set $\{\alpha_k\}$. Hence in general the conjugate diagram will be contained in the smallest closed convex region con-

[17] A. Pringsheim, Sitzungsberichte der Bayer. Akad., 1923.

taining the points $\{\alpha_k\}$. However, we wish to restrict ourselves to the case in which $\bar{\mathfrak{J}}$, the conjugate diagram, is equal to $\mathfrak{J}^*$, the least convex region containing the points α_k and we will therefore impose certain restrictions on this set. Let E denote the set of extreme points of $\mathfrak{J}^*$. Then the restriction we shall impose is that *the set* $E \cap \{\alpha_k\}$ *is everywhere dense on* E, i.e., any neighborhood of a point of E contains an α_k which is in E.

We now show that every point of E is singular for $\gamma(z)$. It is sufficient to show that every point of $E \cap \{\alpha_k\}$ has this property. To prove this let α_ν be a given point of $E \cap \{\alpha_k\}$ and take a supporting line (Stützgerade), $\mathfrak{L}$ of $\mathfrak{J}^*$ going through α_ν.[18] We allow z to approach α_ν along the outer normal to $\mathfrak{L}$; hence, if $d_k = |z - \alpha_k|$, we have $d_k > d_\nu$, $k \neq \nu$, and

$$|(z - \alpha_\nu)\gamma(z) - B_\nu| \leqq \sum_{\substack{k=1 \\ k \neq \nu}}^{\infty} |B_k| \frac{d_\nu}{d_k}.$$

As $d_\nu \to 0$ it is easily seen that the right member can be made arbitrarily small and we therefore conclude that $z = \alpha_\nu$ is a singular point of $\gamma(z)$.

The conjugate diagram $\bar{\mathfrak{J}}$ of $G(z)$ will therefore contain the set E, and, since a closed convex region is determined by its extreme points, will contain $\mathfrak{J}^*$. However, since, as we have seen, $\mathfrak{J}^* \supset \bar{\mathfrak{J}}$, we have $\mathfrak{J}^* = \bar{\mathfrak{J}}$.

We next consider more closely the growth of $G(z)$ in certain directions. Let $re^{i\varphi_1}$ be a given direction such that the supporting line of $\mathfrak{J}$ perpendicular to this direction passes through one and only one point, say $\bar{\alpha}_1$, of the set $\{\bar{\alpha}_\nu\}$. Writing $\bar{\alpha}_\nu = a_\nu e^{i\beta_\nu}$, we have

$$G(re^{i\varphi_1}) = B_1 \exp(a_1 re^{i(\varphi_1 - \beta_1)}) + \sum_{\nu=2}^{\infty} B_\nu \exp(a_\nu re^{i(\varphi_1 - \beta_\nu)}).$$

However, since the projection of $\bar{\alpha}_1$ on the direction $re^{i\varphi_1}$ is, by definition of a supporting line, greater than that of any other $\bar{\alpha}_\nu$, it follows that the first term will dominate all the others so that we have[19]

$$G(z) \sim B_1 e^{\alpha_1 z}, \qquad z = re^{i\varphi_1}. \tag{3}$$

2. **The zeros of $G(z)$ in certain sectors.** The result concerning the growth of $G(z)$ obtained above enables us to get some information concerning its zeros in certain sectors. Let us suppose that the indicator diagram $\mathfrak{J}$ contains as part of its boundary an arc of a smooth curve with a continuously turning tangent. Then the tangents to points of this arc will be supporting lines of $\mathfrak{J}$ corresponding to directions which fill out a certain sector $S: \varphi_1 < \varphi < \varphi_2$. Also each tangent will go through an extreme point of $\mathfrak{J}$ and, since the $\bar{\alpha}_\nu$ lie everywhere dense among the extreme points, there will be a set of angles φ everywhere

[18] For a proof that such an $\mathfrak{L}$ always exists see Carathéodory, Palermo Rendiconti, vol. 32 (1911), p. 198.

[19] This result has also been obtained by Regensburger, loc. cit.

dense in S for which the supporting lines go through points $\bar{\alpha}_\nu$. For these directions we have by (3)

$$(4) \qquad \lim_{r \to \infty} \frac{\log | G(re^{i\varphi}) |}{r} = | \alpha_k | \cos (\beta_k + \varphi).$$

We can now use methods similar to those used in Part I. If we suppose $G(z) \neq 0$ in S, it follows that

$$W(z) = [G(z)]^{-1}$$

is regular and of order unity in S. Moreover, since the limit (4) exists along an everywhere dense set of directions in S, we have $h^*(\varphi) = -h(\varphi)$, where $h(\varphi)$, $h^*(\varphi)$ are the indicators of $G(z)$, $W(z)$, respectively. Taking now any two directions θ, η in S with $| \theta - \eta | < \pi$, we have by the Phragmén-Lindelöf inequality

$$h(\varphi) \leqq H(\varphi) = A \cos \varphi + B \sin \varphi,$$

where $H(\theta) = h(\theta)$, $H(\eta) = h(\eta)$. In the same way

$$h^*(\varphi) = -h(\varphi) \leqq -H(\varphi),$$

so that we deduce

$$h(\varphi) = H(\varphi) = A \cos \varphi + B \sin \varphi.$$

But the curve $\rho = A \cos \varphi + B \sin \varphi$ represents, in polar coördinates, a circle passing through the points $(h(\theta), \theta)$, $(h(\eta), \eta)$ and the origin. Hence, the supporting lines of directions between θ and η must all pass through the point P, where OP is the diameter of the circle, O being the origin. But as we have assumed that the tangents are continuously turning, this cannot be the case. Hence $G(z)$ has at least one zero in S. By a simple argument used previously, we conclude further that there is an infinite number of zeros in S. In fact, by a proof following very closely on the lines of that of Theorem III we can arrive at the following conclusion.

THEOREM IV. *The sum*

$$\sum_{1}^{\infty} | z_\nu |^{-1}$$

taken over the zeros of $G(z)$ in S is divergent.

3. **An example.** Whereas the above results all tend to show that infinite exponential sums have properties similar to those of finite exponential sums, in particular, that the zeros tend to array themselves along the outward drawn normals to the indicator diagram, the following example shows, in contradistinction to the finite case, that exceptions can occur. We show, namely, that a sector may contain an infinite number of zeros and yet contain no normal to the indicator diagram.

Consider the series

$$G(z) = \sum_{\nu=1}^{\infty} B_\nu e^{a_\nu z} \sin b_\nu z,$$

where $b_\nu = \pi/(\nu!)$, $a_\nu = 1 - b_\nu$, and B_ν are real numbers to be described later. It is easily seen that the indicator diagram of this function is a triangle with vertices at the points $(1 - \pi, \pm \pi)$, $(1, 0)$, so that the positive real axis is not an outward drawn normal. Nevertheless, $G(z)$ has an infinite number of zeros along this axis. For let k be an integer and take $z = k!$. Then we have

$$G(k!) = B_{k+1} e^{k! a_{k+1}} \sin \frac{\pi}{k+1} + \sum_{k+2}^{\infty} B_\nu e^{k! a_\nu} \sin \frac{\pi k!}{\nu!}.$$

Hence, we need only choose the B_ν so that

$$|B_{k+1}| > \left(\sum_{k+2}^{\infty} |B_\nu|\right) e^{k!(1-a_{k+1})} \left(\sin \frac{\pi}{k+1}\right)^{-1},$$

and also so that the term

$$B_{k+1} e^{k! a_{k+1}} \sin \frac{\pi}{k+1}$$

has an alternating sign in order to insure the existence of at least one zero between $z = k!$ and $z = (k+1)!$.

PRINCETON UNIVERSITY.

THE HALF-GROUP OF COSETS BELONGING TO A GROUP

By Charles Hopkins

The problem of incorporating into a single system the various quotient-groups associated with a given group G has recently received attention: one finds, for example, a solution in the papers of Ore on structures.[1] As Ore points out, however, the wide applicability of his results is attained "by the elimination of the elements from the algebraic theories". It is the purpose of this paper to present a solution in which the elements of the quotient-groups occupy the center of interest. We shall incorporate the elements of certain quotient-groups associated with G into a multiplicative system, which we call *the half-group belonging to G*. It is not difficult to see that this "multiplicative system" can never be a group if, as seems reasonable, we require that two elements belonging to distinct quotient-groups have a unique product.

I. The half-group $\Gamma(G)$

Let G denote any group containing more than one element, and let Φ denote a set of operators for G, each operator effecting a *proper* automorphism of G. Of the set Φ we require that it contain operators effecting each of the inner isomorphisms[2] of G. Let $H(G)$ denote the set of all subgroups in G which individually admit each operator of Φ. Evidently $H(G)$ is either the set of all normal subgroups of G or a subset of this set. In any case, $H(G)$ will contain both the identity subgroup E and the group G itself. The following are familiar results: if H_a and H_b are any two members of $H(G)$, then the complexes H_aH_b and H_bH_a are identical and each is equal to the union $\{H_a, H_b\}$; both the union and the cross-cut $[H_a, H_b]$ are contained in $H(G)$.

Now each H in $H(G)$ gives rise to the quotient-group $\Gamma = G/H$. Let $Q(G)$ denote the set of *distinct* quotient-groups associated with the set $H(G)$, two quotient-groups Γ_a and Γ_b being regarded as distinct if, and only if, $H_a \neq H_b$. We suppose, furthermore, that two distinct quotient-groups have no element in common. Let Σ denote the set of all group-elements in $Q(G)$; i.e., the logical sum of the sets of elements in Γ_a, Γ_b, etc. We wish to define for the elements of Σ a "multiplication" which shall have the following characteristics:

(1a) the set Σ is closed under multiplication;

(1b) multiplication is associative for any three elements of Σ;

Received June 24, 1937; in revised form, December 14, 1937. I am indebted to Mr. Garrett Birkhoff for suggestions regarding the title and nomenclature of this paper.

[1] *On the foundations of abstract algebra*, I, II, Annals of Mathematics, vol. 36 (1935), pp. 406–437; vol. 37 (1936), pp. 265–292; *Structures and group theory*. I, this Journal, vol. 3 (1937), pp. 149–174.

[2] If G is abelian, the set Φ may be void. Throughout this article we designate simple and multiple "isomorphisms" by the terms *isomorphism* and *homomorphism*, respectively.

(1c) if X and Y are any two elements of Σ, then the product XY is unique;

(1d) for any two elements of the same quotient-group the product "within the set" coincides with the product "within the group".

Now each Γ_a is isomorphic with the group of distinct cosets $H_a x$, where x ranges over the elements of G. Two cosets $H_a x$ and $H_b y$ are certainly distinct if $H_a \neq H_b$, and we have assumed that two distinct quotient-groups Γ_a and Γ_b have no element in common.[3] By setting up an isomorphism θ_a between each Γ_a and the group of cosets $H_a x$ we obtain, therefore, a one-to-one correspondence σ between the distinct cosets of G and the elements of Σ. Following the usual procedure, we define the product $H_a x \cdot H_b y$ to be the coset $H_a H_b xy$. The set of distinct cosets of G will then constitute a multiplicative system having the properties (1a)-(1d) above.

Let $S_{a\alpha}$ and $S_{b\beta}$ denote two elements of Γ_a and Γ_b, respectively, and let $H_a s_\alpha$ and $H_b s_\beta$ denote the cosets to which they correspond under σ; i.e., $(H_a s_\alpha)\theta_a = S_{a\alpha}$, $(H_b s_\beta)\theta_b = S_{b\beta}$. Let us define the product $S_{a\alpha} S_{b\beta}$ by the equation

$$S_{a\alpha} S_{b\beta} = S_{c\gamma}, \tag{1}$$

where

$$S_{c\gamma} = (H_c s_\gamma)\theta_c, \qquad H_c = H_a H_b, \qquad s_\gamma = s_\alpha s_\beta.$$

Under this definition of multiplication within Σ the correspondence σ becomes an isomorphism, and Σ becomes a multiplicatively-closed system satisfying (1a)-(1d) above. This system we shall call the half-group $\Gamma(G)$. In view of the isomorphism σ we may identify $\Gamma(G)$ with the half-groups of cosets of G; in particular, we shall usually designate the elements of G/E—E being the identity of G—by the elements of G itself.

It is evident, of course, that in identifying the elements of Σ with the distinct cosets of G we have merely obtained the most obvious solution of our original problem, which was to incorporate the "abstract" elements of Σ into a multiplicative system satisfying (1a)-(1d) above. Other solutions are always possible: e.g., we can define the product XY to be the product within the quotient-group or the null-quotient G/G, when X and Y belong to the same or to distinct quotient-groups, respectively.

What chiefly distinguishes the half-group $\Gamma(G)$ from an ordinary group is the fact that for two given elements A and C of $\Gamma(G)$ the equation $AY = C$ need not have a solution in $\Gamma(G)$. We shall prove that

For the existence of a solution of the equation

$$S_{a\alpha} Y = S_{c\gamma} \tag{2}$$

it is necessary and sufficient that H_a be contained in H_c; if (2) is satisfied by an element of the quotient-group Γ_y, then the equation

$$X S_{a\alpha} = S_{c\gamma} \tag{2'}$$

is also satisfied by an element of Γ_y, and conversely.

[3] This assumption is not trivial, since in certain connections it is convenient to regard $A/A \cap B$ as a subgroup of K/B, where A and B are two normal subgroups of a group K.

From the definition of multiplication one sees that the condition $H_a \subseteq H_c$ is necessary. To prove its sufficiency we observe that (2) is satisfied by the coset $H_y s_\alpha^{-1} s_\gamma$, belonging to Γ_y, where H_y is any member of $H(G)$ satisfying the equation $H_a H_y = H_c$. (This equation always has at least the solution $H_y = H_c$.) Since $H_a H_y = H_y H_a$, and since (2′) is satisfied by the coset $H_y s_\gamma s_\alpha^{-1}$, it follows that to each solution of (2) contained in Γ_y there corresponds a solution of (2′), and conversely.

Evidently the set of all elements in $\Gamma(G)$ which satisfy equation (2) can be split into subsets, each subset consisting of those solutions which belong to the same quotient-group. Let Γ_b denote a quotient-group which contains at least one solution of (2). The following result is of interest when Γ_b contains only a finite number of solutions of (2):

If the equation $S_{a\alpha} Y = S_{c\gamma}$ (or $X S_{a\alpha} = S_{c\gamma}$) is satisfied by an element of the quotient-group Γ_b, then it is satisfied by exactly λ elements of Γ_b, where λ is the index of H_b in H_c.

For if $S_{bu} (= H_b s_u)$ is a given solution of (2) contained in Γ_b, then any coset $H_b z_a s_u$, where z_a is a variable element of H_a, is also a solution of (2). As z_a ranges over the elements of H_a, we obtain exactly λ distinct cosets (modulo H_b), where λ is the index of the cross-cut $[H_a, H_b]$ in H_a. From a well-known theorem, this index is equal to the index of H_b in $H_a H_b$ $(= H_c)$. That our theorem holds for the equation $X S_{a\alpha} = S_{c\gamma}$ as well follows from the fact that (2) and (2′) have the same number of solutions in Γ_b.

We shall say that H_a is *maximal* in H_c if the equation $H_a H_y = H_c$ has only the one solution $H_y = H_c$. Then for the theorem above we may state as a corollary: *the equation $S_{a\alpha} Y = S_{c\gamma}$, as well as $X S_{a\alpha} = S_{c\gamma}$, will have a unique solution in $\Gamma(G)$ if, and only if, H_a is maximal in H_c.*

In the concluding portion of this section we shall state, without proof, certain results relating to "subsystems" of $\Gamma(G)$, a subsystem being defined as a set of elements in $\Gamma(G)$ which is closed under multiplication. It is easy to see that any subsystem Δ is expressible in one, and only one, way as the direct sum of components Λ_a, Λ_b, etc., where each component Λ_a is either void or is the subgroup consisting of all elements of Δ which occur in Γ_a.

Those elements of $\Gamma(G)$ which are commutative with every element of $\Gamma(G)$ constitute a subsystem, namely, the center of $\Gamma(G)$. One can easily show that the center of $\Gamma(G)$ is the direct sum of the centers of Γ_a, Γ_b, etc.

The cross-cut of two subsystems is always a subsystem or is void. As in the case of ordinary subgroups of a group, the product of two subsystems Δ_x and Δ_y is a subsystem if, and only if, $\Delta_x \Delta_y = \Delta_y \Delta_x$.

Let Δ be any subsystem of $\Gamma(G)$, and let J be a given *set* of elements in $\Gamma(G)$. We shall call Δ a right-hand, or left-hand, or invariant J-subsystem, according as Δ contains ΔJ, or $J\Delta$, or $J\Delta J$. If, in particular, J is the set Σ of all elements in $\Gamma(G)$, then, borrowing the terminology of ring theory, we may call Δ a right-hand, or left-hand, or invariant (two-sided) *ideal*. It is easy to show that every ideal is an invariant ideal and is the direct sum of uniquely determined quotient-groups.

Let $S_{a\alpha}(= H_a s_\alpha)$ be a given element of $\Gamma(G)$, and let φ be a fixed operator in the set Φ. We define the product $S_{a\alpha}\varphi$ to be the element $H_a \cdot s_\alpha \varphi$. Since $H_a\varphi = H_a$ and since each φ effects an automorphism of G, it is evident that Φ is a set of operators for $\Gamma(G)$; i.e., $S_{a\alpha}\varphi \cdot S_{b\beta}\varphi = S_{a\alpha}S_{b\beta}\cdot\varphi$. Any subsystem Δ which admits all the operators in Φ we shall call a *Φ-subsystem*. Evidently $\Delta\Phi \subseteq \Delta$ implies $\Delta\Phi = \Delta$, since each operator effects a proper automorphism of G. As examples of Φ-subsystems, we have the following: the center of $\Gamma(G)$, every quotient-group Γ_a, every subgroup of Γ_a which is homomorphic with a Φ-subgroup of G.

II. The extended group-ring $\mathfrak{R}(G)$

Up to this point we have discussed the half-group $\Gamma(G)$ and its subsystems for groups of unrestricted generality. From this point on we shall assume that G is a denumerable group and that the number of groups in the set $H(G)$ is finite.[4]

Let n denote this number. Then for the members of $H(G)$ there exists at least one sequence such that in this sequence the product of two groups never precedes a factor. We now label these groups so that in the fixed sequence

$$H_1, H_2, \cdots, H_n \tag{1}$$

the subscript k of the product $H_iH_j(= H_k)$ will satisfy the inequalities

$$k \geqq i, \qquad k \geqq j. \tag{2}$$

Corresponding to the sequence (1) we have for the quotient-groups $\Gamma_i\ (=G/H_i)$ the ordered arrangement

$$\Gamma_1, \Gamma_2, \cdots, \Gamma_n. \tag{3}$$

Since $\Gamma_i\Gamma_j = G/H_iH_j$, it is clear that the subscript k of $\Gamma_i\Gamma_j(= \Gamma_k)$ satisfies the inequalities (2).

We know that $H(G)$ contains at least the identity subgroup E and the group G itself;[4] hence, in (1) the first and last terms must be E and G, respectively. Consequently, in (3) the first term is G and the last term is G/G.

Let $\mathbf{F}$ be any commutative field and let $\mathfrak{R}(G)$ be the hypercomplex system over $\mathbf{F}$ whose basis-units are the elements of $\Gamma(G)$. From this point on we shall designate the elements of Γ_i by e_i, u_{i2}, $\cdots$, where e_i is the identity of Γ_i. Employing the usual notation, we then write[5] $\mathfrak{R}(G) = \mathbf{F}e_1 + \mathbf{F}u_{12} + \cdots +$

[4] This does not imply that the number of elements in G is finite. For example, if G is an infinite group whose elements are all of order 2 (E excepted), and if Φ is the set of all proper automorphisms of G, then $H(G)$ contains only the two groups E and G.

The results in this section are valid, for the most part, if we replace "finite" by "denumerable"; in this case, however, the proof of our main theorem may involve an infinite number of constructions.

[5] Thus a given element of $\mathfrak{R}(G)$ is expressible uniquely in the form $\zeta_{11}e_1 + \zeta_{12}u_{12} + \cdots + \zeta_{n1}e_n$, it being understood that only a finite number of the coefficients ζ_{ij} are different from zero when $\mathfrak{R}(G)$ is of infinite rank over $\mathbf{F}$.

$\mathbf{F}e_2 + \mathbf{F}u_{22} + \cdots + \mathbf{F}e_n$. Let $\mathfrak{R}_i$ denote the hypercomplex system over $\mathbf{F}$ whose units are the elements of Γ_i. Then $\mathfrak{R}(G)$ is the direct sum of the subrings $\mathfrak{R}_1, \mathfrak{R}_2, \cdots, \mathfrak{R}_n$. We shall call $\mathfrak{R}(G)$ the *extended* group-ring of G, the terminology being suggested by the fact that each $\mathfrak{R}_i$ is the regular group-ring over $\mathbf{F}$ of the group Γ_i.

One sees readily that the set of n subrings $\mathfrak{R}_1, \cdots, \mathfrak{R}_n$ is closed under multiplication, and that the multiplication table for the $\mathfrak{R}$'s is obtained from that of the Γ's if we replace Γ_i by $\mathfrak{R}_i$. It is evident, therefore, that of these n subrings only one, namely, $\mathfrak{R}_n$, is an *invariant* subalgebra. Our main objective in this section is to show that by an appropriate change of basis we can exhibit $\mathfrak{R}(G)$ as the direct sum of n *invariant* subrings $\bar{\mathfrak{R}}_1, \cdots, \bar{\mathfrak{R}}_i, \cdots, \bar{\mathfrak{R}}_n$, where each $\bar{\mathfrak{R}}_i$ is isomorphic with $\mathfrak{R}_i$ $(i = 1, 2, \cdots, n)$.

We shall prove the following theorem: *$\mathfrak{R}(G)$ is the direct sum of n invariant subrings $\bar{\mathfrak{R}}_1, \cdots, \bar{\mathfrak{R}}_n$; i.e., $\mathfrak{R}(G) = \bar{\mathfrak{R}}_1 + \bar{\mathfrak{R}}_2 + \cdots + \bar{\mathfrak{R}}_n$, where*

(a) $\bar{\mathfrak{R}}_i^2 = \bar{\mathfrak{R}}_i; \bar{\mathfrak{R}}_i\bar{\mathfrak{R}}_j = \bar{\mathfrak{R}}_j\bar{\mathfrak{R}}_i = 0, \qquad i \neq j;$

(b) *$\bar{\mathfrak{R}}_i$ and $\mathfrak{R}_i$ are ring-isomorphic.*

Now $\mathfrak{R}(G)$ contains as a subring the hypercomplex system $\mathfrak{E}(G) = \mathbf{F}e_1 + \mathbf{F}e_2 + \cdots + \mathbf{F}e_n$. We shall first reduce $\mathfrak{E}(G)$ to the direct sum of n fields $\mathbf{F}\bar{e}_i$ which annul each other, and then show that this reduction of $\mathfrak{E}(G)$ leads to the desired reduction of $\mathfrak{R}(G)$.

Since e_i is the identity of Γ_i, the elements $e_1, \cdots, e_n$ constitute the set of all idempotents in $\Gamma(G)$. (Digressing momentarily, we point out that e_1 is the principal unit of $\mathfrak{R}(G)$, while e_n has the properties of a "null-element" in $\Gamma(G)$; i.e., $Xe_n = e_nX = e_n$ for every X in $\Gamma(G)$.) For our purpose the important characteristics of the multiplication table of these idempotents are the following:

$$\text{(4)}\qquad \begin{aligned} &e_i^2 = e_i; \quad e_ie_j = e_je_i = e_k, && (k \geqq i,\ k \geqq j);\\ &e_1e_i = e_i; \quad e_ne_i = e_n && (i, j = 1, 2, \cdots, n). \end{aligned}$$

[See (1), (2), (3) above.]

We now write $\bar{e}_n$ in place of e_n and define $\bar{e}_{n-\mu}$ by the recursive relation

$$\text{(5)}\qquad \bar{e}_{n-\mu} = e_{n-\mu}(e_1 - \bar{e}_{n-\mu+1} - \bar{e}_{n-\mu+2} - \cdots - \bar{e}_{n-1} - \bar{e}_n) \quad (\mu = 1, 2, \cdots, n-1).$$

We shall show that the n elements $\bar{e}_1, \bar{e}_2, \cdots, \bar{e}_n$ satisfy the following conditions:

$$\text{(6)}\qquad \bar{e}_i^2 = \bar{e}_i; \quad \bar{e}_i\bar{e}_j = \bar{e}_j\bar{e}_i = 0 \qquad (i \neq j);$$

$$\text{(7)}\qquad \bar{e}_i = e_i + d_{i,i+1}e_{i+1} + \cdots + d_{in}e_n \qquad (i = 1, 2, \cdots, n),$$

where

(7a) the coefficients d_{ik} are rational integers[6] or zero, and

(7b) d_{ik} is zero whenever the equation $e_iX = e_k$ has no solution in $\Gamma(G)$.

From (4) it is easy to see that (6) is satisfied by $\bar{e}_n$ and $\bar{e}_{n-1}$.

[6] Or congruent to rational integers, in case $\mathbf{F}$ is a modular field.

We suppose, then, that (6) and (7) are known to hold for the elements $\bar{e}_r$, $\bar{e}_{r+1}, \cdots, \bar{e}_n$, where r is a fixed integer greater than 1. We shall prove that these conditions must then hold for $\bar{e}_{r-1}, \bar{e}_r, \cdots, \bar{e}_n$.

Now $\bar{e}_{r-1} = e_{r-1}(e_1 - \bar{e}_r - \bar{e}_{r+1} - \cdots - \bar{e}_n) = e_{r-1} - e_{r-1}\bar{e}_r - \cdots - e_{r-1}\bar{e}_n$, and if we replace $\bar{e}_r$, $\bar{e}_{r+1}$, etc., by their equivalents from (7) and refer to the equations $e_ie_j = e_k$, $k \geqq i$, $k \geqq j$, from (4), then it is clear that $\bar{e}_{r-1}$ is of the form $e_{r-1} + d_{r-1,r}e_r + \cdots + d_{r-1,n}e_n$, where the coefficients $d_{r-1,k}$ are rational integers or zero. Furthermore, in this expression for $\bar{e}_{r-1}$ a term with a non-zero coefficient (e_k, say) can arise only from products of the form $e_{r-1}e_j(= e_k)$. Hence all parts of (7) hold for $\bar{e}_{r-1}$.

Since (6) is assumed to hold for $\bar{e}_r, \cdots, \bar{e}_n$, we obtain for $\bar{e}_{r-1}^2$ the expansion

$$\bar{e}_{r-1}^2 = e_{r-1}^2(e_1 - \bar{e}_r + \cdots + \bar{e}_n - 2\bar{e}_r - \cdots - 2\bar{e}_n) = \bar{e}_{r-1}.$$

Hence $\bar{e}_{r-1}$ is idempotent. Moreover, if $\bar{e}_{r+j}$ is any element of the set $\bar{e}_r, \bar{e}_{r+1}, \cdots, \bar{e}_n$, then

$$\bar{e}_{r-1}\bar{e}_{r+j} = e_{r-1}(e_1 - \bar{e}_r - \cdots - \bar{e}_n)\,\bar{e}_{r+j} = e_{r-1}(e_1\bar{e}_{r+j} - \bar{e}_r\bar{e}_{r+j} - \cdots - \bar{e}_n\bar{e}_{r+j}).$$

Now $e_1\bar{e}_{r+j} = \bar{e}_{r+j}$, and of the remaining products in the last parenthesis exactly one is different from zero, namely, the product $\bar{e}_{r+j}\bar{e}_{r+j}$. Hence $\bar{e}_{r-1}\bar{e}_{r+j} = e_{r-1}(\bar{e}_{r+j} - \bar{e}_{r+j}) = 0$. Thus we have proved by induction that the n elements $\bar{e}_1, \cdots, \bar{e}_n$ satisfy (6) and (7).

Since these n idempotents are linearly independent over $\mathbf{F}$, they can replace $e_1, \cdots, e_n$ as a basis for the ring $\mathfrak{C}(G)$. Each $\mathbf{F}\bar{e}_i$ is a field isomorphic with $\mathbf{F}$, and it is obvious that $\mathfrak{C}(G)$ is the direct sum of these fields. Since each $\mathbf{F}\bar{e}_i$ is an irreducible two-sided (invariant) ideal in $\mathfrak{C}(G)$, it follows from known considerations that the elements $\bar{e}_1, \cdots, \bar{e}_n$ are the *only* n linearly independent idempotents in $\mathfrak{C}(G)$ which satisfy (6) above. If we express the e's in terms of the $\bar{e}$'s, we obtain n equations of the form $e_i = \bar{e}_i + \sum_{t=i+1}^{n} D_t\bar{e}_t$, the coefficients being rational integers or zero. The following equations, which we shall use below, are easily verified:

(8a) $$e_i\bar{e}_i = \bar{e}_i \qquad (i = 1, 2, \cdots, n);$$

(8b) $$e_i\bar{e}_j = 0 \qquad (i > j).$$

We are now in a position to prove the main theorem of this section. In the set of original basis-units for $\mathfrak{R}(G)$, we replace each e_i by $\bar{e}_i$ and each u_{ij} by $\bar{u}_{ij}(= u_{ij}\bar{e}_i)$ $(i = 1, 2, \cdots, n)$, obtaining thereby the set $\bar{e}_1, \bar{u}_{12}, \cdots, \bar{e}_2, \bar{u}_{22}, \cdots, \bar{e}_n$. For the moment we assume that the elements of this set are linearly independent (we shall prove it below); they will then constitute a new set of basis-units for $\mathfrak{R}(G)$. By definition, $\mathfrak{R}_i = \mathbf{F}e_i + \mathbf{F}u_{i2} + \cdots + \mathbf{F}u_{ij} + \cdots$. We denote by $\bar{\mathfrak{R}}_i$ the expression which $\mathfrak{R}_i$ becomes when we replace e_i by $\bar{e}_i$ and each u_{ij} by $\bar{u}_{ij}$ $(j = 2, 3, \cdots)$. From (8a) and from the equation $\bar{u}_{ij} = u_{ij}\bar{e}_i$ we see that $\bar{\mathfrak{R}}_i = \mathfrak{R}_i\bar{e}_i$. Since $\bar{\mathfrak{R}}_i^2 = \mathfrak{R}_i^2\bar{e}_i^2 = \bar{\mathfrak{R}}_i$, it is evident that

$\bar{\mathfrak{R}}_i$ is a subalgebra of $\mathfrak{R}(G)$. Since $\bar{\mathfrak{R}}_i\bar{\mathfrak{R}}_j = \mathfrak{R}_i\bar{e}_i\mathfrak{R}_j\bar{e}_j = \mathfrak{R}_i\mathfrak{R}_j\bar{e}_i\bar{e}_j$, and since $\bar{e}_i\bar{e}_j = 0$ for $i \neq j$, it follows that each $\bar{\mathfrak{R}}_i$ is an *invariant* subalgebra.

To complete the proof of our theorem we have only to show that $\mathfrak{R}_i$ and $\bar{\mathfrak{R}}_i$ are ring-isomorphic. Certainly $\mathfrak{R}_i$ and $\bar{\mathfrak{R}}_i$ are homomorphic under the correspondence defined by $e_i \sim \bar{e}_i$, $u_{ij} \sim \bar{u}_{ij}$, $f \sim f$ for all elements f in $\mathbf{F}$. For if x and y are any two elements of $\mathfrak{R}_i$, and if $\bar{x}$ and $\bar{y}$ are the corresponding elements of $\bar{\mathfrak{R}}_i$, then, since $\bar{x} = x\bar{e}_i$, $\bar{y} = y\bar{e}_i$, we have

$$x + y \sim \overline{x + y} = (x + y)\,\bar{e}_i = \bar{x} + \bar{y};$$

$$xy \sim \overline{xy} = (xy)\bar{e}_i = (x\bar{e}_i)\,(y\bar{e}_i) = \bar{x}\bar{y}.$$

To prove that this homomorphism is an isomorphism it is sufficient to show that $\bar{x}$ is zero only when x is zero. Using (4) and (7a), we obtain the equations

$$\bar{x} = x\bar{e}_i = x(e_i + d_{i,i+1}e_{i+1} + \cdots + d_{in}e_n) = x + x',$$

where x' is a sum of elements from $\mathfrak{R}_{i+1}, \mathfrak{R}_{i+2}, \cdots, \mathfrak{R}_n$. Since $\mathfrak{R}_i$ and $\mathfrak{R}_j$, $i \neq j$, have only the element zero in common, we see that $\bar{x} = 0$ implies $x = 0$. Incidentally, this proves that the elements $\bar{e}_i$, $\bar{u}_{i2}$, $\cdots$ of a given $\bar{\mathfrak{R}}_i$ are linearly independent over $\mathbf{F}$. Let $\bar{z}$ be any linear function, with coefficients in $\mathbf{F}$, of the elements $\bar{e}_1$, $\bar{u}_{12}$, $\cdots$, $\bar{e}_2$, $\bar{u}_{22}$, $\cdots$, $\bar{e}_n$, and let $\bar{z}_1$, $\bar{z}_2$, $\cdots$, $\bar{z}_n$ be the components of $\bar{z}$ which lie in $\bar{\mathfrak{R}}_1, \cdots, \bar{\mathfrak{R}}_n$, respectively. Assume that $\bar{z} = 0$ and that $\bar{z}_a \neq 0$. If $\bar{z} = \bar{z}_1 + \cdots + \bar{z}_a + \cdots + \bar{z}_n = 0$, then $\bar{z}\bar{e}_a = \bar{z}_a\bar{e}_a = \bar{z}_a = 0$. Hence, the elements $\bar{e}_1, \cdots, \bar{u}_{ij}, \cdots, \bar{e}_n$ are linearly independent, since we have already established the linear independence of the subset which occurs in a given $\bar{\mathfrak{R}}_i$.

A decomposition of a hypercomplex system $\mathfrak{S}$ into the direct sum of invariant subrings is usually designated by the term *direct decomposition* of $\mathfrak{S}$. The decomposition $\bar{D}$ of $\mathfrak{R}(G)$ into the direct sum of the components $\bar{\mathfrak{R}}_i$ above is probably the most useful direct decomposition of $\mathfrak{R}(G)$, but it is obvious that $\mathfrak{R}(G)$ will in general have direct decompositions other than $\bar{D}$. Even if we know of a direct decomposition D that its components, in some order, are ring-isomorphic with the components of $\bar{D}$, it does not always follow that D and $\bar{D}$ are the same decomposition. [The simplest example arises by choosing for G the non-cyclic group of order 4, for Φ the identical automorphism, and for $\mathbf{F}$ the field of rational numbers.] We mention a striking exception:

If G is a group G_p in which the order of every element is a power of a given prime p, and $\mathbf{F}$ *is a commutative field* $\mathbf{F}_p$ *of characteristic p, then*

(i) *each component $\bar{\mathfrak{R}}_i$ of $\bar{D}$ above is direct-indecomposable*;

(ii) *$\bar{D}$ is the only direct decomposition of $\mathfrak{R}(G)$ into direct-indecomposable components.*

We sketch the proof of (i). Each $\mathfrak{R}_i$ in $\mathfrak{R}(G_p)$ is the group-ring of a p-group Γ_i over $\mathbf{F}_p$; a group-ring of this sort is a primary ring—i.e., every divisor of zero is nilpotent—and contains, accordingly, no idempotents other than 0

and the principal unit. Since $\bar{\mathfrak{R}}_i$ and $\mathfrak{R}_i$ are ring-isomorphic, $\bar{\mathfrak{R}}_i$ contains no idempotents except 0 and $\bar{e}_i$. Since a direct decomposition of $\bar{\mathfrak{R}}_i$ would involve a decomposition of $\bar{e}_i$ into the sum of two or more idempotents each different from zero, it follows that $\bar{\mathfrak{R}}_i$ is direct-indecomposable.

Proof of (ii).[7] Let $\mathfrak{S}$ be any hypercomplex system with a principal unit e, and let D' and D'' be two direct decompositions of $\mathfrak{S}$ which have as components $\mathfrak{S}_1', \cdots, \mathfrak{S}_r'$ and $\mathfrak{S}_1'', \cdots, \mathfrak{S}_s''$, respectively. Let e_i' and e_j'' denote the components of e in $\mathfrak{S}_i'$ and $\mathfrak{S}_j''$, respectively. Now each e_i' (and each e_j'') is idempotent, in the center of $\mathfrak{S}$, and different from zero. Since $\mathfrak{S}_i'$ and $\mathfrak{S}_j''$ are invariant subrings of $\mathfrak{S}$, we know that $e_i'e_j''$ is an idempotent element e_{ij} contained in both $\mathfrak{S}_i'$ and $\mathfrak{S}_j''$. Hence those e_{ij}'s which are not zero must be distinct; furthermore, the product of two distinct e_{ij}'s is necessarily zero. Since

$$\mathfrak{S} = \mathfrak{S}e = \mathfrak{S}e^2 = \mathfrak{S}(e_1' + \cdots + e_r')(e_1'' + \cdots + e_s''),$$

it follows that $\mathfrak{S}$ is the direct sum $\sum \mathfrak{S}e_{ij}$, where the summation extends over those e_{ij}'s which are not zero. Each component $\mathfrak{S}e_{ij}$ is an invariant subring of $\mathfrak{S}$; consequently, the decomposition $\mathfrak{S} = \sum \mathfrak{S}e_{ij}$ is a direct decomposition D'''. Since

$$\mathfrak{S}_i' = \mathfrak{S}e_i' = \mathfrak{S}e_i'e = \mathfrak{S}(e_{i1}' + \cdots + e_{is}') = \sum_t \mathfrak{S}e_{it},$$

we see that by a proper grouping of the components of D''' we recover the components of D' (and of D'', as well). And it is evident that D''' will be a refinement of at least one of the original decompositions if, and only if, these original decompositions D' and D'' are distinct.

Returning now to $\mathfrak{R}(G_p)$, we choose for D'' the decomposition $\bar{D}$ and for D' any second direct decomposition of $\mathfrak{R}(G_p)$ in which each component is direct-indecomposable. Since each component of $\bar{D}$ is direct-indecomposable (see (i)), the decomposition D''' above must coincide with $\bar{D}$ and with D', for it can be a refinement of neither. Therefore D' and $\bar{D}$ must be the same direct decomposition of $\mathfrak{R}(G_p)$. This completes the proof of (ii).

We touch briefly upon the significance of $\bar{D}$ in connection with the two regular representations of $\Gamma(G)$ when G is a finite group. It is known that any representation of a finite group over a commutative field can be derived from a representation of its group-ring, and conversely—a result which can be extended to the half-group $\Gamma(G)$. From the properties of $\bar{D}$—in particular, from the fact that $\bar{\mathfrak{R}}_i$ and $\mathfrak{R}_i$ are ring-isomorphic—it follows that the first (or second) regular representation of $\Gamma(G)$ is equivalent under a linear transformation T to the direct sum of n representations $M_1, M_2, \cdots, M_n$, where each M_i is equivalent to the first (or second) regular representation of Γ_i. From (7a)

[7] Although each $\bar{\mathfrak{R}}_i$ is direct-indecomposable, it is never irreducible for $i \neq n$. Hence the familiar "uniqueness-theorem" for semi-simple rings is not available.

it is clear that T may be chosen so that its coefficients are rational integers. More generally, *any* representation of $\Gamma(G)$ by linear transformations over **F**, whether regular or not, is equivalent to the direct sum of representations of the quotient-groups $\Gamma_1, \cdots, \Gamma_n$. Since each Γ_i is a homomorphic image of $G(=\Gamma_1)$, the theory of representations of $\Gamma(G)$ is contained essentially in the theory of representations for G. For example, if the characteristic of **F** is prime to the order of G, then any representation of $\Gamma(G)$ over **F** is completely reducible, and to obtain an extension of **F** in which this representation splits into its absolutely-irreducible components we need only to choose an extension of **F** in which the components belonging to G have this property.

TULANE UNIVERSITY.

SPHEROIDAL AND BIPOLAR COÖRDINATES

By H. Bateman

1. **The relations between the different coördinates.** Let $x = w \cos \phi$, $y = w \sin \phi$, then if

$$z = r\cos\theta = kuv = kS \operatorname{sh} \sigma, \qquad u \geqq 1, \qquad -1 \leqq v \leqq 1, \tag{1.1}$$

$$w = r \sin \theta = k(u^2 - 1)^{\frac{1}{2}}(1 - v^2)^{\frac{1}{2}} = kS \sin \tau, \tag{1.2}$$

$$k(u - v) = R, \qquad k(u + v) = R', \qquad S(\operatorname{ch} \sigma - \cos \tau) = 1, \tag{1.3}$$

$$(u - v)e^{\sigma} = u + v, \qquad (u^2 - v^2)\cos \tau = u^2 + v^2 - 2, \tag{1.4}$$

$$r^2 = k^2(u^2 + v^2 - 1). \tag{1.5}$$

It is usual to call (r, θ, ϕ) the spherical polar coördinates, (z, w, ϕ) the cylindrical coördinates, (u, v, ϕ) the spheroidal coördinates and (σ, τ, ϕ) the bipolar coördinates of the point P whose rectangular coördinates are (x, y, z).

For a second point P_0 whose rectangular coördinates are (x_0, y_0, z_0), quantities $u_0, v_0, w_0, \theta_0, \phi_0, \sigma_0, \tau_0, R_0, R_0', S_0$ may be defined by similar equations with a constant k_0 which may or may not be different from k. We shall, however, be interested in a function $G(x, y, z, x_0, y_0, z_0)$ which is harmonic when considered as a function of x, y, z and also when considered as a function of x_0, y_0, z_0. For reasons of symmetry it will be convenient in this case to take $k_0 = k$.

2. **The standard spheroidal harmonics.** It is well known that Laplace's equation has the simple solutions

$$P_n^m(u)P_n^m(v)e^{im\phi}, \qquad Q_n^m(u)P_n^m(v)e^{im\phi},$$

where $P_n^m(u)$ and $Q_n^m(u)$ are associated Legendre functions.

In the case of symmetry about the axis of z the simple solutions become

$$P_n(u)P_n(v) \quad \text{and} \quad Q_n(u)P_n(v).$$

A series of solutions of the second type is particularly useful for the representation of a potential function in the space outside a prolate spheroid whose foci are at the points with rectangular coördinates $(0, 0, k)$, $(0, 0, -k)$, respectively. This leads to the consideration of Neumann series of type

$$f(u) = \sum_{n=0}^{\infty} (2n + 1)c_n Q_n(u). \tag{2.1}$$

Such a series is known to converge in the region of the complex u-plane that lies outside an ellipse with the points $+k$ and $-k$ as foci, when $f(u)$ is an analytic

Received July 25, 1937 and August 27, 1937; in revised form October 30, 1937.

function with no singularities outside this ellipse. Useful expressions for the coefficients which supplement those given by Neumann may be obtained by using certain polynomials expressible as generalized hypergeometric functions

$$F_n(z) = F(-n, n+1, \tfrac{1}{2} + \tfrac{1}{2}z; 1, 1; 1), \tag{2.2}$$

$$Z_n(z) = F(-n, n+1; 1, 1; z), \tag{2.3}$$

the notation being the usual one except that the suffixes adopted by Barnes have been dropped, the semicolons being sufficient to specify the number of parameters in the numerator and the denominator of the function in the present work.

These polynomials are such that if D_x denotes the operator d/dx,

$$\begin{aligned} F_n(D_x) \operatorname{cosech} x &= \operatorname{cosech} x\, P_n(\coth x), \\ F_n(D_x) x \operatorname{cosech} x &= \operatorname{cosech} x\, Q_n(\coth x), \\ Z_n(ze^{-x}) &= e^{ze^{-x}} F_n(2D_x - 1)e^{-ze^{-x}}, \\ Z_n(-D_x)x^{-1} &= x^{-1}P_n(1 - 2x^{-1}). \end{aligned} \tag{2.4}$$

Moreover the coefficient c_n is given by the formulas[1]

$$c_n = (-)^n \lim_{x \to 0} F_n(D_x) \operatorname{cosech} x\, f(\coth x), \tag{2.5}$$

$$c_n = \lim_{x \to 0} Z_n(-D_x) 2x^{-1} f(1 + 2x^{-1}), \tag{2.6}$$

which may be applied, for example, to the function $f(z) = 1/(z - t)$. Let us now write $u_0 = \coth \rho$, $W^2 = u_0^2 + u^2 + v^2 - 1 + 2uu_0v$, $WT = uu_0 + v$ and apply the formula (2.5) to the two well-known expansions

$$W^{-1} = \sum_{n=0}^{\infty} (-)^n (2n+1) P_n(u) P_n(v) Q_n(u_0), \tag{2.7}$$

$$W^{-1} Q_0(T) = \sum_{n=0}^{\infty} (-)^n (2n+1) Q_n(u) P_n(v) Q_n(u_0), \qquad (u_0 > u). \tag{2.8}$$

Since

$$W^2 \operatorname{sh}^2\rho (\operatorname{ch} \sigma - \cos \tau) = \operatorname{ch}(\sigma + 2\rho) - \cos \tau,$$

we obtain the following representations of spheroidal harmonics

$$\begin{aligned} P_n(u)P_n(v) &= S^{-1}F_n(2D_\sigma)S, \qquad S = (\operatorname{ch} \sigma - \cos \tau)^{-\frac{1}{2}}, \\ Q_n(u)P_n(v) &= S^{-1}F_n(2D_\sigma)[SQ_0(u)]. \end{aligned} \tag{2.9}$$

[1] See H. Bateman, Annals of Mathematics, vol. 35 (1934), pp. 767–775 for the first of these. The second, (2.6), is new.

The series (2.8) may not be so well known as (2.7), but it is easily summed by expressing it as a definite integral

$$\tfrac{1}{2}\int_{-1}^{1}\frac{dt}{u_0 - t}\,[(t + uv)^2 + (u^2 - 1)(1 - v^2)]^{-\frac{1}{2}},$$

which has the value indicated.

The more general series

$$\sum_{n=0}^{\infty}(2n + 1)Q_n(u_0)P_n(v_0)Q_n(u)P_n(v)$$

represents a function harmonic in both (x, y, z) and (x_0, y_0, z_0), which, when $v = v_0 = 1$, is represented by the series

$$V = \sum_{n=0}^{\infty}(2n + 1)Q_n(u_0)Q_n(u) = [Q_0(u_0) - Q_0(u)](u - u_0)^{-1}$$

$$= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(m + n + 1)^{-1}u_0^{-m-1}u^{-n-1},$$

where the summation extends over all combinations of values for which $m + n$ is an even positive integer or zero.

When $r > k$ and $r_0 > k$ the expansion of V in a series of Legendre functions is, with the same restriction on m and n,

$$V = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{k}{r_0}\right)^{m+1}\left(\frac{k}{r}\right)^{n+1}P_m(\cos\theta_0)P_n(\cos\theta)(m + n + 1)^{-1}.$$

In particular, when $v_0 = 1$ and consequently $\theta_0 = 0$, we have the expansion

$$\sum_{n=0}^{\infty}(2n + 1)Q_n(u)P_n(v)Q_n(u_0) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left(\frac{k}{r}\right)^{n+1}u_0^{-m-1}P_n(\cos\theta)(m + n + 1)^{-1}.$$

3. **The standard bipolar harmonics.** The standard bipolar harmonics are

$$(u - v)^{-1}e^{\nu\sigma}P_\nu^m(\cos\tau)e^{im\phi} \quad\text{and}\quad (u - v)^{-1}e^{\nu\sigma}Q_\nu^m(\cos\tau)e^{im\phi},$$

where ν and m are arbitrary constants. Taking $m = 0$ to get the case of symmetry about the axis of z, we shall show that

$$(u - v)^{-1}e^{\nu\sigma}P_\nu(\cos\tau) = \sum_{n=0}^{\infty}(2n + 1)F_n(-2\nu - 1)Q_n(u)P_n(v).$$

Denoting the left side by V, we remark that V is a potential function which, when regarded as a function of u, has singularities only when $u^2 = v^2$ and when $\cos\tau \leqq -1$, the latter giving a singularity only when ν is not an integer. Now $\cos\tau \leqq -1$ for $v^2 < u^2 \leqq 1$. Hence, if $-1 \leqq v \leqq 1$ and $u > 1$, the series is convergent whatever be the value of ν. For the left side can then be expanded in a convergent Neumann Q-series whose coefficients may be found by our rule

by first putting $v = 1$. When $\nu = 0$ the series reduces to Heine's series for $(u - v)^{-1}$; and when $\nu = -\frac{1}{2}$ it gives the expansion

$$(2\pi^{-1})(u^2 - v^2)^{-\frac{1}{2}}K[(1 - v^2)^{\frac{1}{2}}(u^2 - v^2)^{-\frac{1}{2}}] = \sum_{n=0}^{\infty} (4n + 1)(-\tfrac{1}{2}/, n)^2 Q_{2n}(u)P_{2n}(v),$$

where $K(k)$ is the quarter period of the Jacobi elliptic functions with modulus k and $(z/, n)$ is used to denote the coefficient of t^n in the binomial expansion of $(1 + t)^z$, while $(n, /z)$ may be used for the reciprocal of this coefficient.

The potential function V can also be expanded in a series of spherical harmonics in the region outside the sphere $r = k$. The axial value of V is in fact

$$\left(\frac{k}{z-k}\right)\left(\frac{z+k}{z-k}\right)^{\nu} = \frac{1}{2}\left(\frac{z+k}{z-k}\right)^{\nu+1} - \frac{1}{2}\left(\frac{z+k}{z-k}\right)^{\nu}$$
$$= \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{k}{z}\right)^n [g_n(\nu + 1) - g_n(\nu)],$$

where $g_n(x)$ is the polynomial of Mittag-Leffler.[2] Hence, if $r > k$,

$$V = \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{k}{r}\right)^n P_{n-1}(\cos\theta)[g_n(\nu + 1) - g_n(\nu)].$$

When $n > 0$ the polynomial $g_n(x)$ can be expressed in terms of the hypergeometric function; indeed

$$g_n(x) = 2xF(1 - n, 1 - x; 2; 2),$$

while $g_0(x) = 1$.

At a point on the sphere $r = k$ we have $u^2 + v^2 = 2$ and so

$$V_s = (u + v)^{\nu}(u - v)^{-\nu-1}P_{\nu}(0).$$

For some values of ν this function can be represented by the Legendre series

$$V_s = \frac{1}{2}\sum_{n=1}^{\infty} P_{n-1}(\cos\theta)[g_n(\nu + 1) - g_n(\nu)].$$

Noting that $u + v = 2\cos\frac{1}{2}\theta$, $u - v = 2\sin\frac{1}{2}\theta$, we have

$$V = \tfrac{1}{4}\cos^{\nu}(\tfrac{1}{2}\theta)\operatorname{cosec}^{\nu+1}(\tfrac{1}{2}\theta)\,P_{\nu}(0) = 2^{-\frac{1}{2}}(1 + \mu)^{\frac{1}{2}\nu}(1 - \mu)^{-\frac{1}{2}\nu-\frac{1}{2}}P_{\nu}(0),$$

where $\mu = \cos\theta$. By the known theory of expansions in series of Legendre functions the expansion of this function is permissible when $-3 < 2\nu < 1$.

Spherical potential problems in spheroidal coördinates

4. **Values of the harmonics on the sphere $r = k$.** At a point on the sphere $r = k$ the spheroidal coördinates u, v are connected by the relation $u^2 + v^2 = 2$

[2] G. Mittag-Leffler, Acta Mathematica, vol. 15 (1891), pp. 1–32; vol. 29 (1905), pp. 101–181.

and we may put $z = k \tanh t$, $w = k \operatorname{sech} t$. Then, if $s = e^{-t}$, we have at this point on the sphere

$$(4.1)\qquad \begin{aligned} u &= (1 + s)(1 + s^2)^{-\frac{1}{2}} = a, \qquad \text{say},\\ v &= (1 - s)(1 + s^2)^{-\frac{1}{2}} = c, \qquad \text{say},\end{aligned}$$

and the formulas for the standard spheroidal harmonics of symmetric type become

$$(4.2)\qquad \begin{aligned} (1 + s^2)^{-\frac{1}{2}} P_n(a)P_n(c) &= F_n(2D_t - 1)(1 + s^2)^{-\frac{1}{2}},\\ (1 + s^2)^{-\frac{1}{2}} Q_n(a)P_n(c) &= F_n(2D_t - 1)[(1 + s^2)^{-\frac{1}{2}} Q_0(u)].\end{aligned}$$

The first of these equations indicates that when $| s | < 1$ we have the expansion

$$(4.3)\qquad (1 + s^2)^{-\frac{1}{2}} P_n(a)P_n(c) = \sum_{m=0}^{\infty} (-\tfrac{1}{2}/, m) F_n(-4m - 1)s^{2m},$$

where $(z/, m)$ is used for the coefficient of t^m in the expansion of $(1 + t)^z$, and $(m, /z)$ for the reciprocal of this coefficient.

If we make the substitutions

$$(4.4)\qquad \begin{aligned} P_{2n}(a)P_{2n}(c) &= \sum_{p=0}^{n} (-)^{n+p}(n + p - \tfrac{1}{2}/, n)(n/, p)P_{2p}(b),\\ P_{2n+1}(a)P_{2n+1}(c) &= \sum_{p=0}^{n} (-)^{n+p}(n + p + \tfrac{1}{2}/, n)(n/, p)P_{2p+1}(b),\end{aligned}$$

where $b = (1 - s^2)/(1 + s^2)$, and make use of the expansion

$$(4.5)\qquad (1 + s^2)^{-p} P_q(b) = \sum_{m=0}^{\infty} s^{2m}(-p/, m)F(-q, q + 1, -m; p, 1; 1),$$

which may be new, we are led to the identities

$$(4.6)\qquad F_{2n}(-4m - 1) = \sum_{p=0}^{n} (-)^{n+p}(n + p - \tfrac{1}{2}/, n)(n/, p)F_{2p}^{-\frac{1}{2}}(-2m - \tfrac{1}{2}),$$

$$(4.7)\qquad F_{2n+1}(-4m - 1) = \sum_{p=0}^{n} (-)^{n+p}(n + p + \tfrac{1}{2}/, n)(n/, p)F_{2p+1}^{-\frac{1}{2}}(-2m - \tfrac{1}{2}),$$

wherein use has been made of Pasternack's notation

$$(4.8)\qquad F_n^m(z) = F(-n, n + 1, \tfrac{1}{2} + \tfrac{1}{2}z + \tfrac{1}{2}m; 1, m + 1; 1).$$

It may be remarked that there is a second expansion associated with (4.5), viz.,

$$(4.9)\qquad \begin{aligned} &(1 + s^2)^{-p} P_q(-b)\\ &\qquad = \sum_{n=0}^{\infty} (n + p - 1/, n)(-s^2)^n F(-q, q + 1, p + n; p, 1; 1).\end{aligned}$$

5. **Expansions in series of spheroidal harmonics.** When the value of a potential function V is known on the sphere $r = k$ and V is symmetric about the axis of z, the expansions

$$V = \sum_{n=0}^{\infty} a_n P_n(u) P_n(v), \tag{5.1}$$

$$V = \sum_{n=0}^{\infty} c_n Q_n(u) P_n(v), \tag{5.2}$$

when they exist, may, perhaps, be obtained by first expressing V by the well-known integral which gives the solution of the Dirichlet problem for the sphere. This integral is

$$\pm V = \frac{1}{4\pi} \int_0^{\pi} \sin\theta' \, d\theta' \int_0^{2\pi} d\phi' k(k^2 - r^2)\rho^{-3} V', \tag{5.3}$$

where V' is the value of V at the point with spherical polar coördinates (k, θ', ϕ'), $\rho^2 = k^2 + r^2 - 2kr(\cos\theta \cos\theta' + \sin\theta \sin\theta' \cos(\phi - \phi'))$, and the upper or lower sign is taken according as r is less than or greater than k.

To obtain the coefficients in the expansions (5.1) and (5.2) we put $u = 1$ in the first and $v = 1$ in the second and in each case expand the resulting axial value of V, namely, V_a. For this we need the expansions

$$(u^2 - 1)[(u-1)^2 + t^2(u+1)^2]^{-\frac{3}{2}} = \sum_{n=0}^{\infty} (2n+1)(-)^n U_n^*(t) Q_n(u), \tag{5.4}$$

$$(1 - v^2)[(1+v)^2 + T^2(1-v)^2]^{-\frac{3}{2}} = \sum_{n=0}^{\infty} (n + \tfrac{1}{2}) V_n^*(T) P_n(v), \tag{5.5}$$

where $t = \tan\frac{1}{2}\theta'$, $T = \cot\frac{1}{2}\theta'$.

Regarding the first of these as a Neumann Q-series and using our rule for the determination of the coefficients, we have the formula

$$U_n^*(e^{2x}) = F_n(D_x + 3) U_0^*(e^{2x}), \tag{5.6}$$

where

$$U_0^*(t) = (1 + t^2)^{-\frac{3}{2}}.$$

Since $(u-1)^2 + t^2(u+1)^2$ is zero when $u = e^{i\theta'}$, the Q-series may be expected to converge outside an ellipse in the u-plane which has the points ± 1 as foci and which passes through the point $u = e^{i\theta'}$. This ellipse meets the real axis where $u = (1 + \sin^2\theta')^{\frac{1}{2}}$. Now the greatest value of $\sin^2\theta'$ is 1, and so the Q-series certainly converges when $u > 2^{\frac{1}{2}}$. When this condition is satisfied, the Q-series for V_a is

$$V_a = \sum_{n=0}^{\infty} (4n+2)(-)^n Q_n(u) \int_0^{\infty} (1 + t^2)^{-\frac{1}{2}} t \, dt U_n^*(t) g(t), \tag{5.7}$$

where $V' = g(\tan \frac{1}{2}\theta')$. The coefficients c_n are thus found. With a suitable type of function $g(t)$ the series for V_a may converge for some values of u that are smaller than $2^{\frac{1}{2}}$.

In the particular case in which $V = Q_m(u)P_m(v)$ we have $g(t) = Q_m(a)P_m(c)$, where

$$a = \frac{t+1}{(t^2+1)^{\frac{1}{2}}}, \qquad c = \frac{t-1}{(t^2+1)^{\frac{1}{2}}}, \tag{5.8}$$

and the formula for V_a indicates that

$$\begin{aligned} \int_0^\infty (1+t^2)^{-\frac{1}{2}} Q_m(a)P_m(c)U_n^*(t)t\,dt &= 0 \qquad (m \neq n) \\ &= (4n+2)^{-1} \qquad (m = n). \end{aligned} \tag{5.9}$$

This result has been checked in some particular cases.

6. The function $V_n^*(t)$ may be defined in a similar way by the equation

$$V_n^*(e^{2x}) = F_n(D_x + 3)V_0^*(e^{2x}), \tag{6.1}$$

where

$$tV_0^*(t) = \int_0^\infty \frac{s\,ds}{s+t}(1+s^2)^{-\frac{1}{2}}. \tag{6.2}$$

Putting $s = e^{2y}$, we may write

$$e^{3x}V_0^*(e^{2x}) = \int_{-\infty}^\infty (\tfrac{1}{2}\operatorname{sech} 2y)^{\frac{1}{2}} \operatorname{sech}(x-y)\,dy.$$

Hence, since $F_n(D_x)\operatorname{sech}(x-y) = \operatorname{sech}(x-y)P_n[\tanh(x-y)]$, we have

$$e^{3x}V_n^*(e^{2x}) = \int_{-\infty}^\infty (\tfrac{1}{2}\operatorname{sech} 2y)^{\frac{1}{2}} \operatorname{sech}(x-y)P_n[\tanh(x-y)]\,dy.$$

That is,

$$\begin{aligned} tV_n^*(t) &= 2\int_{-\infty}^\infty \frac{e^{4y}\,dy}{e^{2x}+e^{2y}}(1+e^{4y})^{-\frac{1}{2}}P_n\left(\frac{e^{2x}-e^{2y}}{e^{2x}+e^{2y}}\right) \\ &= \int_0^\infty \tanh a \operatorname{sech} a \frac{da}{t+\operatorname{sh} a} P_n[(t-\operatorname{sh} a)/(t+\operatorname{sh} a)]. \end{aligned}$$

Making the substitution $v(t + \operatorname{sh} a) = t - \operatorname{sh} a$, we obtain the formula

$$V_n^*(t) = \int_{-1}^1 (1-v^2)P_n(v)\,dv[(1+v)^2 + t^2(1-v)^2]^{-\frac{1}{2}}, \tag{6.3}$$

which gives rise to the equation (5.5). If, moreover, we write (6.2) in the form

$$e^{3x}V_0^*(e^{2x}) = \int_{-\infty}^\infty \frac{2\,dy}{1+e^{-2y}} e^{3x}U_0^*(e^{2x+2y}), \tag{6.2$'$}$$

and operate on both sides with $F_n(D_x)$, we obtain the equation

$$(6.4)\qquad tV_n^*(t) = \int_0^\infty \frac{s\,ds}{s+t}\,U_n^*(s).$$

The expansion (5.5) is convergent for $-1 < v < 1$ when $T > 0$, and we infer, if $V' = f(T) = f(\cot \frac{1}{2}\theta')$, that

$$(6.5)\qquad V_a = \sum_{n=0}^{\infty} (2n+1)P_n(v) \int_0^\infty (1+T^2)^{-\frac{1}{2}} V_n^*(T) f(T) T\,dT.$$

The coefficients a_n are thus determined.

In the particular case in which $V = P_m(u)P_m(v)$ we have $f(T) = P_m(a)P_m(c)$, where

$$(6.6)\qquad a = (1+T)(1+T^2)^{-\frac{1}{2}}, \qquad c = (1-T)(1+T^2)^{-\frac{1}{2}},$$

and the formula indicates that

$$(6.7)\qquad \begin{aligned} \int_0^\infty (1+T^2)^{-\frac{1}{2}} P_m(a)P_m(c)V_n^*(T)T\,dT &= 0 \qquad (m \neq n), \\ &= (2n+1)^{-1} \qquad (m = n). \end{aligned}$$

This result has been checked in some special cases.

It should be noticed that when $c_n = a_n$ and $r = k$ there is a relation between the values of the potential functions represented by the series (5.1) and (5.2).

Let us start with the equation

$$(6.8)\qquad t(t^2+1)^{-\frac{1}{2}}Q_0(a) = \tfrac{1}{2}\int_0^\infty \frac{t\,ds}{s+t}(1+s^2)^{-\frac{1}{2}} \qquad (t > 0),$$

in which a is given by (5.8). The equation is easily proved by making the substitution $s = \operatorname{sh} z$.

Writing $t = e^x$ and operating with $F_n(2D_x - 1)$, we find that, if c is also given by (5.8),

$$\begin{aligned} (1+e^{-2x})^{-\frac{1}{2}}Q_n(a)P_n(c) &= \tfrac{1}{2}\int_0^\infty (1+s^2)^{-\frac{1}{2}}\,ds F_n(2D_x - 1)(1+se^{-x})^{-1} \\ &= \tfrac{1}{2}\int_{-\infty}^\infty (1+e^{-2y})^{-\frac{1}{2}}\,dy\, e^{-y} F_n(2D_y - 1)(1+e^{-x-y})^{-1} \\ &= \tfrac{1}{2}\int_{-\infty}^\infty (1+e^{-2y})^{-\frac{1}{2}}\,dy F_n(2D_y + 1)(e^y + e^{-x})^{-1}. \end{aligned}$$

Using the operation of functional integration by parts which depends on the properties of adjoint differential operators with constant coefficients, we find that

$$\begin{aligned} (1+e^{-2x})^{-\frac{1}{2}}Q_n(a)P_n(c) &= \tfrac{1}{2}\int_{-\infty}^\infty (e^y + e^{-x})^{-1}\,dy F_n(-2D_y + 1)(1+e^{-2y})^{-\frac{1}{2}} \\ &= \tfrac{1}{2}\int_{-\infty}^\infty (e^y + e^{-x})^{-1}\,dy (1+e^{-2y})^{-\frac{1}{2}} P_n(a')P_n(c'), \end{aligned}$$

where

$$a' = (e^{-y} + 1)(e^{-2y} + 1)^{-\frac{1}{2}}, \qquad c' = (e^{-y} - 1)(e^{-2y} + 1)^{-\frac{1}{2}}.$$

Hence

$$(t^2 + 1)^{-\frac{1}{2}} Q_n(a) P_n(c) = \tfrac{1}{2} \int_0^\infty \frac{1}{s + t} (s + 1)^{-\frac{1}{2}} P_n(a') P_n(c')\, ds, \tag{6.9}$$

where

$$a = (t + 1)(t^2 + 1)^{-\frac{1}{2}}, \qquad c = (t - 1)(t^2 + 1)^{-\frac{1}{2}},$$
$$a' = (s + 1)(s^2 + 1)^{-\frac{1}{2}}, \qquad c' = (s - 1)(s^2 + 1)^{-\frac{1}{2}}.$$

This equation shows that the potential function

$$V = \int_0^\infty [(z + s)^2 + w^2]^{-\frac{1}{2}} (s^2 + 1)^{-\frac{1}{2}} P_n(a') P_n(c')\, ds$$

has for $z > 0$ the axial value $V_a = 2(z^2 + 1)^{-\frac{1}{2}} Q_n(\bar{a}) P_n(\bar{c})$, where

$$\bar{a} = (z + 1)(z^2 + 1)^{-\frac{1}{2}}, \qquad \bar{c} = (z - 1)(z^2 + 1)^{-\frac{1}{2}}.$$

The analysis of §§5 and 6 may be useful for the determination of a distribution of axial sources which give rise to a potential V.

1°. *When the sources cover only a finite portion of the z-axis and V is known at points of a very large sphere which contains all the sources.* The solution of the problem is given formally by the series (5.7) when use is made of Neumann's equation

$$Q_n(u) = \tfrac{1}{2} \int_{-1}^{1} \frac{P_n(v)\, dv}{u - v}.$$

2°. *When the sources lie on the portion of the z-axis for which $z < 0$ and V is known on the portion of this axis for which $z > 0$.* Use can then be made of equation (6.9) and related series of products of Legendre functions. The orthogonal relations (5.9) and (6.7) may then be useful for the determination of the coefficients in such expansions.

7. **Definite integrals for the Legendre functions.** When there is only a single source on the axis of z, our analysis leads to expressions for the Legendre functions. One of these will now be derived by an independent method.

Starting from the equation

$$\frac{1}{z + 1} = \int_0^\infty \frac{t\, dt}{(z^2 + t^2)^{\frac{1}{2}}(1 + t^2)^{\frac{1}{2}}} \qquad (z > 0), \tag{7.1}$$

and putting $t = zu$, $z = e^{2x}$, we get

$$\tfrac{1}{2} \operatorname{sech} x = e^{3x} \int_0^\infty u\, du (1 + u^2)^{-\frac{1}{2}} U_0^*(ue^{2x}). \tag{7.2}$$

Operating on both sides with $F_n(D_x)$, we obtain the equation

$$\tfrac{1}{2}\operatorname{sech} x\, P_n(\tanh x) = e^{3x}\int_0^\infty u\,du(1+u^2)^{-\frac{1}{2}}U_n^*(ue^{2x}), \tag{7.3}$$

which is readily transformed into

$$\frac{1}{z+1}P_n\left(\frac{z-1}{z+1}\right) = \int_0^\infty \frac{U_n^*(t)t\,dt}{(z^2+t^2)^{\frac{1}{2}}} \qquad (z>0). \tag{7.4}$$

This equation may be useful for the solution of the integral equation

$$f(z) = \int_0^\infty (z^2+t^2)^{-\frac{1}{2}}g(t)t\,dt \qquad (z>0),$$

which occurs in potential theory when an attempt is made to find a potential of a simple layer over the plane $z = 0$ producing an assigned value on the axis of z above the plane of a potential symmetric about the axis of z. The method suggested by the equation is, indeed, to expand $f(z)$ in a Legendre series of type

$$f(z) = \frac{1}{z+1}\sum_{n=0}^\infty a_n P_n\left(\frac{z-1}{z+1}\right),$$

and to represent $g(t)$, if possible, by a corresponding series

$$g(t) = \sum_{n=0}^\infty a_n U_n^*(t).$$

A "Fourier rule" for this type of expansion is suggested by the orthogonal relation (5.9).

The relation (7.4) may also be obtained by considering the potential function

$$V = \int_0^\infty e^{-zt}J_0(wt)U_n(t)\,dt,$$

where $U_n(t) = e^{-t}Z_n(t)$. It is known that when $w = 0$ and $z > 0$ this potential function has the axial value

$$V = \frac{1}{z+1}P_n\left(\frac{z-1}{z+1}\right).$$

The rule for finding the potential of a simple layer on the plane $z = 0$ which will give rise to the potential V in the space $z > 0$ then gives the equation

$$U_n^*(t) = \int_0^\infty J_0(xt)U_n(x)x\,dx,$$

if (7.4) is assumed to be correct. An independent proof of the last equation can, however, be obtained by starting with the equation

$$U_0^*(t) = \int_0^\infty e^{-x}J_0(xt)x\,dx = t^{-2}\int_0^\infty e^{-z/t}J_0(z)z\,dz,$$

i.e.,

$$e^{4x}U_0^*(e^{2x}) = \int_0^\infty U_0(ze^{-2x})J_0(z)z\,dz.$$

Operating on both sides with $F_n(D_x - 1)$ and using the equations

$$F_n(D_x + 3)U_0^*(e^{2x}) = U_n^*(e^{2x}), \qquad F_n(D_x - 1)U_0(ze^{-2x}) = U_n(ze^{-2x}),$$

we obtain the equation

$$e^{4x}U_n^*(e^{2x}) = \int_0^\infty U_n(ze^{-2x})J_0(z)z\,dz,$$

which is merely another form of the desired equation.

It may be remarked that when $z = 0$ we have $V = (1 + w^2)^{-\frac{1}{2}}P_n(a)P_n(c)$, where

$$a = (w + 1)(w^2 + 1)^{-\frac{1}{2}}, \qquad c = (w - 1)(w^2 + 1)^{-\frac{1}{2}}.$$

A derivation of V from its values on the plane $z = 0$ also leads to the formula

$$\frac{1}{z+1}\, P_n\left(\frac{z-1}{z+1}\right) = \int_0^\infty \frac{zt\,dt}{(z^2 + t^2)^{\frac{3}{2}}}\,(1 + t^2)^{-\frac{1}{2}}P_n(a)P_n(c),$$

where now

$$a = (t + 1)(t^2 + 1)^{-\frac{1}{2}}, \qquad c = (t - 1)(t^2 + 1)^{-\frac{1}{2}}.$$

A consideration of the potential

$$V = \int_0^\infty e^{-zt}J_0(wt)V_n(t)\,dt,$$

in which $V_n(t)$ is the function considered in former papers,[3] likewise leads to the formulas

$$V_n^*(t) = \int_0^\infty J_0(xt)V_n(x)x\,dx,$$

$$\frac{2}{z-1}\, Q_n\left(\frac{z+1}{z-1}\right) = \int_0^\infty \frac{V_n^*(t)t\,dt}{(z^2 + t^2)^{\frac{1}{2}}},$$

$$2(1 + w^2)^{-\frac{1}{2}}Q_n(a)P_n(c) = \int_0^\infty J_0(wt)V_n(t)\,dt.$$

In deriving these results it is helpful to consider a potential which arises from a distribution of charge on the negative z-axis, the density at distance z from the origin being $(1 + z)^{-1}P_n[(z - 1)/(z + 1)]$.

The preceding results may also be derived by writing the first potential in the form

$$Z_n(-D_z)[(z + 1)^2 + w^2]^{-\frac{1}{2}},$$

which indicates that it arises from a set of charges concentrated at the point $z = -1$ on the negative z-axis.

[3] H. Bateman, this Journal, vol. 2 (1936), pp. 569–577; Annals of Mathematics, vol. 38 (1937), pp. 303–310.

The orthogonal relations involving the functions $U_n^*(w)$, $V_n^*(w)$ may be derived from physical considerations in which the mutual energy of two electrostatic fields is calculated in different ways. It is convenient then to place the line charge giving rise to the second type of potential on the positive z-axis instead of the negative. One formula for the mutual energy then takes the form of an integral involving two Legendre functions which is zero when the orders are different. Other formulas giving the desired orthogonal relations are obtained by applying Green's theorem in various ways to the half-space bounded by the plane $z = 0$. A further use of our potentials may be found by expressing the standard bipolar harmonic in the form

$$k^{-1}(u - v)^m(u + v)^{-m-1}P_m\left(\frac{u^2 + v^2 - 2}{u^2 - v^2}\right) = \int_0^\infty e^{-zt-tk}J_0(wt)F(t)\,dt \quad (z > k),$$

where $F(t) = F(-m; 1; 2t) = L_m(2t)$ when m is a positive integer. Since

$$Z_n(t) = \sum_{m=0}^{n} a_{n,m}L_m(2t),$$

where

$$P_n(z) = \sum_{m=0}^{n} a_{n,m}z^m,$$

we find on putting $z = 0$ (a legitimate process because, when $m + 1$ is a positive integer, the above representation of the bipolar harmonic is valid for $z > -k$) that we obtain the expansion

$$P_n(a)P_n(c) = \sum_{m=0}^{n} a_{n,m}P_m[(w^2 - 1)/(w^2 + 1)],$$

which has already been used. This expansion is a particular case of a more general one given on p. 395 of my *Partial Differential Equations of Mathematical Physics* and derived independently by W. N. Bailey, Proceedings of the London Mathematical Society, (2), vol. 41 (1936), pp. 215–220.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

DIFFERENTIABLE AND RIEMANN METRIC

By S. Bochner

In a previous paper[1] the author has proved that an n-dimensional compact analytic space S_n can be mapped topologically-analytically onto the Euclidean E_{2n+1} provided S_n has an analytic Riemann metric.

At first sight the notion of Riemann metric appears to be a very special and "arbitrary" case of the general concept of metric as introduced into topology by Fréchet and Hausdorff, and thus, one might think, our theorem makes a topological conclusion depend on an assumption falling well outside the domain of topology. In the present note we shall discuss the mutual relation of the two metrics for coördinate spaces in general. The discussion will be very simple indeed, but it will show a possibility of characterizing the Riemann metric by properties as closely topological as the situation permits.

1. **A new characterization of Riemann metric.** In a (sufficiently small) n-dimensional coördinate neighborhood S_n of class C_2 with a positive definite or semi-definite tensor $g_{ij}(x)$ of class C_1, the length function

$$L(C) = \int_C \sqrt{g_{ij}(x)\,\dot{x}_i\dot{x}_j}\,dt \tag{1}$$

(for curves of class C_1) gives rise to a geodesic distance $R(x, y)$ having the following properties of a distance function:

1. $D(x, y) \geqq 0$,
2. $D(x, x) = 0$,
3. $D(x, y) + D(y, z) \geqq D(x, z)$,
4. $D(x, y) = D(y, x)$.

If g_{ij} is definite throughout or becomes semi-definite in isolated points only, we have the further property

5. $D(x, y) > 0$ if $x \neq y$.

Also, if S_n belongs to C_p, $p \geqq 2$, and g_{ij} to C_{p-1}, then the square of the geodesic distance

$$\Omega(x, y) = R(x, y)^2 \tag{2}$$

Received September 24, 1937.

[1] *Analytic mapping of compact Riemann spaces into Euclidean space*, this Journal, vol. 3(1937), pp. 339–354.

belongs to C_{p-2}, as a function of the $2n$ variables on the space $S_n \times S_n$. If S_n and g_{ij} are analytic, $\Omega(x, y)$ is also analytic.

Conversely, *if S_n belongs to C_p, $p \geqq 2$, and if a distance function $D(x, y)$ on S_n satisfies properties* 1, 2, 3 *and has the further property that its square*

$$\Omega(x, y) = D(x, y)^2 \tag{3}$$

belongs to C_p on $S_n \times S_n$, then the length function $L(C)$ corresponding to $D(x, y)$ can be represented in the form (1), *the generating tensor g_{ij} belonging to class C_{p-2}. If S_n and $\Omega(x, y)$ are analytic, g_{ij} is also analytic.* In this statement, $L(C)$ is defined by the relation

$$L(C) = \text{l.u.b.} \left\{ \sum_{\kappa=1}^{k-1} D(x^\kappa, x^{\kappa+1}) \right\}, \tag{4}$$

where $x^1, \cdots, x^k$ is any set of successive points on C.

The proof is quite easy. If we replace the variables y_i by new variables $\eta_i = y_i - x_i$, $\Omega(x, y)$ goes over into a new function $W(x, \eta)$ and by Taylor's formula,

$$W(x, \eta) = W(x, 0) + g_i(x)\eta_i + g_{ij}(x)\eta_i\eta_j + \epsilon_{ij}(x, \eta)\eta_i\eta_j,$$

where

$$g_i(x) = \frac{\partial W(x, 0)}{\partial \eta_i}, \qquad g_{ij}(x) = \frac{\partial^2 W(x, 0)}{\partial \eta_i \partial \eta_j} \tag{5}$$

and

$$\epsilon_{ij}(x, \eta) \to 0 \textit{ as } \eta \to 0, \textit{ uniformly in } x. \tag{6}$$

By property 2, $W(x, 0) = 0$, and by property 1, $g_i(x) = 0$, hence

$$W(x, \eta) = g_{ij}(x)\eta_i\eta_j + \epsilon_{ij}(x, \eta)\eta_i\eta_j. \tag{7}$$

By (5), g_{ij} belongs to C_{p-2}, and it is analytic, if $\Omega(x, y)$ is analytic. The tensor character of $g_{ij}(x)$ follows easily from the scalar character of $W(x, \eta)$. The integral representation (1) follows from (7) and (6) in much the same way as in the Euclidean case, the only complication arising from the fact that the norm $N(x)$ of the quadratic form

$$g_{ij}(x)\zeta_i\zeta_j \tag{8}$$

may vanish in points of C. The complication can be easily handled by choosing a finite number of intervals on C along which $N(x) \leqq 2\epsilon$, and outside of which $N(x) \geqq \epsilon$, and by showing that the contribution arising from these intervals to both the expressions (1) and (4) becomes arbitrarily small with ϵ.

If a length function $L(C)$ with the properties

(α) $L(C) \geqq 0$,

(β) $L(0) = 0$,

(γ) $L(C_1) + L(C_2) = L(C_1 + C_2)$

can be generated from a distance function $D(x, y)$ by relation (4), then the generating function is in general not unique. One of the distance functions is uniquely determined, namely, the geodesic distance $R(x, y)$ which is defined as the greatest lower bound of $L(C)$ for all curves C joining x, y. Obviously $R(x, y)$ is characterized by the property that for any other generating $D(x, y)$,

$$D(x, y) \leqq R(x, y).$$

The existence of distance functions other than $R(x, y)$ is well known. For instance, we may choose

$$D(x, y) = R(x, y)[1 - R(x, y)^2]^{\frac{1}{2}}. \tag{9}$$

It is obvious that properties 1, 2, 4, 5 carry over from $R(x, y)$ to $D(x, y)$. As for property 3, all we have to prove is that for sufficiently small positive numbers A, B, C, the relation $A + B \geqq C$ implies the relation

$$A(1 - A^2)^{\frac{1}{2}} + B(1 - B^2)^{\frac{1}{2}} \geqq C(1 - C^2)^{\frac{1}{2}}$$

and this can be verified by a simple calculation.

If we put

$$R(x, y)^2 \equiv W(x, \eta) = g_{ij}(x)\eta_i\eta_j + g_{ijk}(x)\eta_i\eta_j\eta_k + g_{ijkl}(x)\eta_i\eta_j\eta_k\eta_l + \cdots, \tag{10}$$

the square of (9) differs from (10) in the values of the coefficients g_{ijkl}. Thus two distance functions generating the same length (1) may differ in the coefficients of the fourth powers in η of the corresponding functions $W(x, \eta)$; whereas the coefficients g_{ij}, if symmetric in their indices, must be the same. Now, the coefficients g_{ijk} are also uniquely determined, as a consequence of property 4. In fact, this property implies the relation

$$W(x, \eta) = W(x + \eta, -\eta),$$

or

$$g_{ij}(x)\eta_i\eta_j + g_{ijk}(x)\eta_i\eta_j\eta_k + \cdots = \left(g_{ij}(x) + \frac{\partial g_{ij}}{\partial x_k}\eta_k + \cdots\right)\eta_i\eta_j - (g_{ijk}(x) + \cdots)\eta_i\eta_j\eta_k + \cdots,$$

and therefore

$$\left(\frac{1}{6}\frac{\partial^3 W(x, 0)}{\partial\eta_i\partial\eta_j\partial\eta_k}\right) \equiv g_{ijk} = \frac{1}{2}\frac{\partial g_{ij}}{\partial x_k}.$$

2. Points of contraction on analytic Riemann spaces. The points x at which (8) is semi-definite are those at which the determinant

$$| g_{ij}(x) | \tag{11}$$

has the value 0. They are characterized by the property that there are small curves C_1 issuing from x whose Riemann length becomes small of higher order

than the Euclidean length. We shall term them "points of contraction". On non-analytic Riemann spaces property 5 is compatible with the existence of points of contraction having points of accumulation in the interior of the space. For instance, if $n = 1$, property 5 holds for $D(x, y) = | \varphi(x) - \varphi(y) |$, provided $\varphi(x)$ is a strictly monotone function of the given class of differentiability, without extra restrictions on the set of points at which its derivative vanishes.

But for an analytic space, if $D(x, y)^2$ is an analytic function on $S_n \times S_n$, the corresponding points of contraction are isolated. In fact, the function (11) being analytic, the set of points on which it vanishes is, in every closed neighborhood, a finite sum of analytic cells of different dimensions.[2] If there were a cell of dimension $\geqq 1$, there would exist an analytic curve C whose length (1) had the value 0. The geodesic distance between the endpoints of this curve would also be 0, in contradiction to property 5.

If S_n is compact, the points of contraction, if any, are finite in number. If there are none, our mapping theorem holds. One should expect the theorem to remain true in the general case, possibly with the qualification that the mapping functions need not be analytic at the points of contraction themselves. The author was unable to settle the question. We only mention, without giving details, that a direct generalization of the original method allows one to establish the theorem under additional restrictions on the behavior of $D(x, y)$ in the neighborhood of the critical points. The restrictions are rather severe. They are elaborations of the requirement that, the origin being a point of contraction, the quantity $D(x, y)$ dominate and be dominated by functions of the form

$$[\max (x, y)]^\lambda \cdot [(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2]^{\frac{1}{2}},$$

where λ is a positive integer.

PRINCETON UNIVERSITY.

[2] S. Lefschetz, *Topology*, 1930, Chapter VIII.

LINEAR FUNCTIONALS SATISFYING PRESCRIBED CONDITIONS

By Ralph Palmer Agnew

1. **Introduction.** A function (or transformation) $q \equiv q(x)$ with domain and range in linear spaces is called *linear* if

$$q(ax + by) = aq(x) + bq(y) \qquad (a, b \in R;\ x, y \in E), \tag{1.01}$$

where E is the domain of q and R is the set of real numbers. If the range of $q(x)$ is in R, then $q(x)$ is called a *functional*. Using notation of Banach[1] we call a functional $p(x)$ a *p-function* if

$$p(tx) = tp(x) \qquad (t \geqq 0;\ x \in E), \tag{1.02}$$

$$p(x + y) \leqq p(x) + p(y) \qquad (x, y \in E). \tag{1.03}$$

We denote the class of linear functionals $f \equiv f(x)$ by F and the class of p-functions by P.

A theorem of Banach (loc. cit., p. 29) of which we make repeated use is

Theorem 1.1. *If $p \in P$, then there exists $f \in F$ with*

$$f(x) \leqq p(x) \qquad (x \in E). \tag{1.11}$$

Since each linear functional f is also a p-function, i.e., $F \subset P$, the following theorem, of which we shall make explicit and implicit use, is trivial.

Theorem 1.2. *If $f \in F$, then there exists $p \in P$ with $f(x) \leqq p(x)$ for all $x \in E$.*

Let $p_0 \in P$ and a set Ψ of pairs $\{x, y\}$ of elements $x, y \in E$ be prescribed. One problem in which we shall be interested is that of determining whether there exist linear functionals $f \in F$ possessing the properties

$$f(x) \leqq p_0(x) \qquad (x \in E), \tag{1.21}$$

$$f(y) = f(x) \qquad (\{x, y\} \in \Psi). \tag{1.22}$$

We assume Ψ has the property that if $\{x, y\} \in \Psi$ then $\{y, x\} \in \Psi$, and that $\{x, x\} \in \Psi$ for each $x \in E$; this assumption is convenient and entails no loss of generality.

We shall say that a p-function $p \equiv p(x)$ *enforces* a specified property (or set of properties) if every $f \in F$, with $f(x) \leqq p(x)$ for all $x \in E$, must possess the specified property (or set of properties).

For example, a slight amplification of work of Banach (loc. cit., p. 33) shows

Received September 27, 1937; presented to the American Mathematical Society, October 30, 1937.

[1] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 28.

that if E is the space of real bounded functions $x \equiv x(s)$ defined over $-\infty < s < \infty$, then

$$p_B(x) = \operatorname*{g.l.b.}_{n>0;\, \lambda_k \in R} \ \varlimsup_{s\to\infty} \frac{1}{n} \sum_{k=1}^{n} x(s + \lambda_k) \tag{1.23}$$

is a p-function which enforces the properties

$$f(x) \leqq \varlimsup_{s\to\infty} x(s) \qquad (x \in E), \tag{1.24}$$

and

$$f(x(s+\lambda)) = f(x(s)) \qquad (\lambda \in R;\ x \in E). \tag{1.25}$$

The interest in $f(x)$ lies in the fact that $\operatorname*{Lim}_{s\to\infty} x(s) \equiv f(x)$ is a generalization of $\lim_{s\to\infty} x(s)$ which exists for all real bounded functions. The rôle of the analogue (1.24) of (1.21) is to ensure that the generalized limit of $x(s)$ lies between the inferior and superior limits of $x(s)$; and the rôle of the analogue (1.25) of (1.22) is to ensure that the generalized limits of $x(s)$ and $x(s + \lambda)$ are equal. This example and related ones will be discussed in §9.

There is of course no à priori reason for believing that there exists $p \in P$ which enforces specified properties. The situation is governed by

THEOREM 1.3. *In order that there exist $f \in F$ having a specified set of properties, it is necessary and sufficient that there exist at least one $p \in P$ which enforces these properties.*

Proof of this theorem is quite trivial. To prove sufficiency, choose $p_1 \in P$ which enforces the properties. By Banach's Theorem 1.1, there exists $f_1 \in F$ with $f_1(x) \leqq p(x)$ and hence this f_1 must have the properties. To prove necessity, let $f_1 \in F$ have the properties. Then f_1 itself is a p-function which enforces the properties; for if $f(x)$ is a linear functional with $f(x) \leqq f_1(x)$ for all $x \in E$, then $-f(x) = f(-x) \leqq f_1(-x) = -f_1(x)$, so $f_1(x) \leqq f(x)$ for all $x \in E$. Hence $f(x) = f_1(x)$, and $f(x)$ has the properties in question. This proves the theorem.

The preceding definitions and theorems suggest the main problem of this paper, namely, that of characterizing analytically the class of p-functions which enforce a specified property or set of properties. The part of the paper from §4 onward deals largely with problems of this type. §§2 and 3 give lemmas involving p-functions and r-functions needed in later sections.

It is known[2] that, if G is a solvable group as in §8 and $p_0 \in P$ is such that $p_0(g(x)) = p_0 x$ for all $g \in G$, $x \in E$, there exists a linear functional f with the properties

$$f(x) \leqq p_0(x) \qquad (x \in E), \tag{1.41}$$

$$f(g(x)) = f(x) \qquad (g \in G;\ x \in E). \tag{1.42}$$

[2] R. P. Agnew and A. P. Morse, *Extensions of linear functionals, with applications to limits, integrals, measures, and densities*, Theorem 3. This paper is to appear in the Annals of Mathematics.

Here Ψ is the set of all pairs $\{x, g(x)\}$ obtained by taking $g \epsilon G$, $x \epsilon E$. In §8 we give a specific p-function which enforces (1.41) and (1.42), and characterizes the class of p-functions which enforce these properties. §§9 and 10 give applications to limits and integrals.

2. **Properties of p-functions.** In this section, we give as lemmas some properties of p-functions of which explicit and implicit use will be made later.

LEMMA 2.1. *In order that a functional $p(x)$ may be a p-function, it is necessary and sufficient that*

$$p(tx) = tp(x), \qquad p(x + y) \leqq p(x) + p(y) \qquad (t > 0; x, y \epsilon E). \tag{2.11}$$

Necessity is obvious from (1.02) and (1.03). To prove sufficiency, we require, in addition to (2.11), only the property

$$p(0) = 0; \tag{2.12}$$

and this follows on putting $x = 0$ and $t = 2$ in the first formula of (2.11).

LEMMA 2.2. *If $p \epsilon P$ and $x_0 \epsilon E$, then a necessary and sufficient condition that*

$$p(x + x_0) = p(x) \qquad (x \epsilon E) \tag{2.21}$$

is that $p(\pm x_0) = 0$, i.e., $p(+x_0) = 0$ and $p(-x_0) = 0$.

To prove necessity, suppose (2.21) holds. Using (2.11), (2.12) and the results obtained by setting $x = -x_0$ and $x = x_0$ in (2.21), we find $p(\pm x_0) = 0$. To prove sufficiency, suppose $p(\pm x_0) = 0$. Then use of (2.11) gives

$$p(x + x_0) \leqq p(x) + p(x_0) = p(x),$$

and

$$p(x) = p[(x + x_0) - x_0] \leqq p(x + x_0) + p(-x_0) = p(x + x_0),$$

from which (2.21) follows.

LEMMA 2.3. *If $p \epsilon P$, then the set E_0 of $x \epsilon E$ for which $p(\pm x) = 0$ forms a linear manifold in E.*

This is easily proved with the aid of Lemma 2.2. The set of $x \epsilon E$ for which $p(x) = 0$ does not ordinarily form a linear manifold in E.

LEMMA 2.4. *If $p \epsilon P$; $x_1, x_2 \epsilon E$; and $p[\pm(x_2 - x_1)] = 0$, then $p(x_2) = p(x_1)$.*

This follows from Lemma 2.2 since it justifies writing

$$p(x_1) = p[x_1 + (x_2 - x_1)] = p(x_2).$$

LEMMA 2.41. *If $p \epsilon P$; $x_1, x_2 \epsilon E$; and $p(x_2) = p(x_1)$, then $p[\pm(x_2 - x_1)] \geqq 0$.*

From $p(x_2) = p[(x_2 - x_1) + x_1] \leqq p(x_2 - x_1) + p(x_1)$ and our hypothesis follows $p(x_2 - x_1) \geqq 0$. Likewise $p(x_1 - x_2) \geqq 0$ and Lemma 2.41 is proved.

LEMMA 2.5. *If $p \epsilon P$, then*

$$-p(-x) \leqq p(x) \text{ and } -p(x) \leqq p(-x) \qquad (x \epsilon E). \tag{2.51}$$

The conclusions in (2.51) are obtained by transpositions in the inequality

$$0 = p(0) = p(x - x) \leqq p(x) + p(-x) \qquad (x \in E). \tag{2.52}$$

It is easy to show that the reverse inequality $p(x) \leqq -p(-x)$ cannot hold for all $x \in E$ unless $p(x)$ is linear.

LEMMA 2.6. *If $p \in P$ and, for some $x_0 \in E$, $p(\pm x_0) \leqq 0$, then $p(\pm x_0) = 0$.*

LEMMA 2.7. *If p_1, $p_2 \in P$ and*

$$p_1(x) \leqq p_2(x) \qquad (x \in E), \tag{2.71}$$

then

$$-p_2(-x) \leqq -p_1(-x) \leqq p_1(x) \leqq p_2(x) \qquad (x \in E). \tag{2.72}$$

LEMMA 2.8. *If $f \in F$, $p \in P$, and $f(x) \leqq p(x)$ for all $x \in E$, then*

$$-p(-x) \leqq f(x) \leqq p(x) \qquad (x \in E). \tag{2.81}$$

LEMMA 2.9. *If $p \in P$ and γ is a linear transformation with domain and range in E, then the functional defined by*

$$p_1(x) = p(\gamma(x)) \qquad (x \in E) \tag{2.91}$$

is a p-function.

Since $F \subset P$, Lemma 2.8 is a corollary of Lemma 2.7. Proofs of Lemmas 2.6, 2.7, and 2.9 are left to the reader.

3. **Properties of r-functions.** We now give, for future reference, the definition of r-functions and theorems involving them.[3] A functional $r(x)$ is called an *r-function* if there exists $f \in F$ with $f(x) \leqq r(x)$ for all $x \in E$. In (3.11) and (3.12) below, $\sum x_k$ stands for the sum $x_1 + \cdots + x_n$ of elements $x_k \in E$.

THEOREM 3.1. *In order that a functional $r(x)$ defined over E may be an r-function, it is necessary and sufficient that*

$$\operatorname*{g.l.b.}_{n, t_k > 0;\, \Sigma x_k = 0} \sum_{k=1}^{n} \frac{r(t_k x_k)}{t_k} \geqq 0. \tag{3.11}$$

THEOREM 3.2. *If $r(x)$ is an r-function, then the functional $p^{(r)}(x)$ defined by*

$$p^{(r)}(x) = \operatorname*{g.l.b.}_{n, t_k > 0;\, \Sigma x_k = x} \sum_{k=1}^{n} \frac{r(t_k x_k)}{t_k} \tag{3.21}$$

is a p-function with

$$-r(-x) \leqq -p^{(r)}(-x) \leqq p^{(r)}(x) \leqq r(x);$$

moreover, if $p \in P$ and $p(x) \leqq r(x)$ over E, then

$$-p^{(r)}(x) \leqq -p(-x) \leqq p(x) \leqq p^{(r)}(x) \qquad (x \in E).$$

[3] R. P. Agnew, *On existence of linear functionals defined over linear spaces.* This paper is to appear in the Bulletin of the American Mathematical Society.

THEOREM 3.3. *In order that $p \in P$ may have the property $p(x) \leqq r(x)$ for all $x \in E$, it is necessary and sufficient that $p(x) \leqq p^{(r)}(x)$ for all $x \in E$.*

The last theorem follows easily from Theorem 3.2. It is a consequence of Banach's Theorem 1.1 that each p-function is an r-function.

In §7 we use the following theorem which was not stated in the paper cited above, but which follows from a slight modification of work in the paper.

THEOREM 3.4. *Let $r(x)$ be a functional defined over E and let $p^{(r)}(x)$ be defined by (3.21). If $p^{(r)}(x)$ is finite for at least one $x \in E$, then $r(x)$ is an r-function and $p^{(r)}(x)$ is a p-function with $p^{(r)}(x) \leqq r(x)$.*

4. **Characterization of p-functions having specified properties.** In the following and some later theorems, it is possible to replace $p_0 \in P$ by a functional $r(x)$ not necessarily a p-function and thereby obtain more general theorems. However, existence of $f \in F$ with $f(x) \leqq r(x)$ is, by Theorem 3.3, equivalent to existence of $f \in F$ with $f(x) \leqq p^{(r)}(x)$; hence we confine ourselves here to p-functions.

THEOREM 4.1. *Let $p_0 \in P$ be fixed. In order that $p \in P$ may enforce the property*

$$f(x) \leqq p_0(x) \qquad (x \in E), \tag{4.11}$$

it is necessary and sufficient that

$$p(x) \leqq p_0(x) \qquad (x \in E). \tag{4.12}$$

Sufficiency is obvious from our definitions; for, if $f(x) \leqq p(x)$ and (4.12) holds, then (4.11) will hold. To establish necessity, let $p \in P$ and suppose $x_0 \in E$ exists such that $p(x_0) > p_0(x_0)$. Let E_0 be the linear manifold in E consisting of elements of the form ax_0 with $a \in R$, and let f_0 be the linear functional defined over E_0 by the formula $f_0(ax_0) = ap(x_0)$. If $a \geqq 0$, then $f_0(ax_0) = p(ax_0)$; while if $a < 0$, then $f(ax_0) = ap(x_0) = -p(\mid a \mid x_0) \leqq p(-\mid a \mid x_0) = p(ax_0)$. Thus we have $f_0(x) \leqq p(x)$ for all $x \in E_0$. Therefore Banach's theorem (loc. cit., pp. 27–28) on extension of linear functionals furnishes a functional $f \in F$ such that $f(x) \leqq p(x)$ for all $x \in E$ and $f(x) = f_0(x)$ for all $x \in E_0$. In particular, $f(x_0) = f_0(x_0) = p(x_0) > p_0(x_0)$. Thus (4.11) fails and necessity of (4.12) follows.

DEFINITION 4.13. If $p_0 \in P$ enforces a set S of properties, if each $p \in P$ with $p(x) \leqq p_0(x)$ for all $x \in E$ enforces S, and if each $p \in P$ with $p(x) > p_0(x)$ for at least one $x \in E$ fails to enforce S, then we shall call p_0 the *greatest p-function which enforces S.*

It is clear that if the greatest $p \in P$ which enforces S exists, it is unique. In this terminology, Theorem 4.1 states that p_0 is the greatest $p \in P$ which enforces (4.11).

THEOREM 4.2. *In order that $p_2 \in P$ may enforce all properties which $p_1 \in P$ enforces, it is necessary and sufficient that*

$$p_2(x) \leqq p_1(x) \qquad (x \in E). \tag{4.21}$$

Necessity follows from Theorem 4.1; and sufficiency is a consequence of the fact that if (4.21) holds, then the class of $f \in F$ with $f(x) \leqq p_2(x)$ is a subclass of

the class of $f \epsilon F$ with $f(x) \leqq p_1(x)$. It follows from Theorem 4.2 that two p-functions p_1, $p_2 \epsilon P$ enforce the same properties only when they are identical,

When properties which some p-functions enforce (or do not enforce) are known, Theorem 4.2 furnishes a *comparison test* useful for determination of properties which other p-functions enforce (or do not enforce).

Let Ψ be (as in §1) a set of pairs $\{x, y\}$ of elements $x, y \epsilon E$ having the property that if $\{x, y\} \epsilon \Psi$ then $\{y, x\} \epsilon \Psi$ and that $\{x, x\} \epsilon \Psi$ for each $x \epsilon E$.

THEOREM 4.3. *In order that $p \epsilon P$ may enforce the property*

$$f(y) = f(x) \qquad (\{x, y\} \epsilon \Psi), \tag{4.31}$$

it is necessary and sufficient that

$$p[\pm(y - x)] = 0 \qquad (\{x, y\} \epsilon \Psi). \tag{4.32}$$

It is easy to see that Theorem 4.3 follows from the following theorem involving a single element $x_0 \epsilon E$.

THEOREM 4.33. *In order that $p \epsilon P$ may enforce the property*

$$f(x_0) = 0, \tag{4.34}$$

it is necessary and sufficient that

$$p(\pm x_0) = 0. \tag{4.35}$$

We prove Theorem 4.33. Sufficiency of (4.35) follows from the inequality

$$-p(-x_0) \leqq f(x_0) \leqq p(x_0).$$

To prove necessity, suppose either $p(x_0) \neq 0$ or $p(-x_0) \neq 0$. Let $\xi_0 = \pm x_0$ according as $p(\pm x_0) \neq 0$. As in the proof of Theorem 4.1, there exists $f \epsilon F$ with $f(x) \leqq p(x)$ for all $x \epsilon E$ and $f(\xi_0) = p(\xi_0)$. Thus $f(x_0) = \pm f(\pm x_0) = \pm f(\xi_0) = \pm p(\xi_0) \neq 0$; hence (4.34) fails and necessity of (4.35) follows. This proves Theorem 4.33 and therefore Theorem 4.3. On account of our definition of Ψ, Theorem 4.3 remains true if we remove the negative sign in (4.32).

THEOREM 4.4. *Let $p_0 \epsilon P$ and Ψ be fixed. In order that $p \epsilon P$ may enforce the two properties*

$$f(x) \leqq p_0(x) \qquad (x \epsilon E), \tag{4.41}$$

$$f(y) = f(x) \qquad (\{x, y\} \epsilon \Psi), \tag{4.42}$$

each of the four conditions

$$p(x) \leqq Q_A(x) \equiv \underset{n>0;\, a_k \epsilon R;\, \{x_k, y_k\} \epsilon \Psi}{\text{g.l.b.}} p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right], \tag{4.43}$$

$$p(x) \leqq Q_B(x) \equiv \underset{n>0;\, \{x_k, y_k\} \epsilon \Psi}{\text{g.l.b.}} p_0\left[x + \frac{1}{n} \sum_{k=1}^{n} (y_k - x_k)\right], \tag{4.44}$$

$$p(x) \leqq Q_C(x) \equiv \underset{n>0;\, \{x_k, y_k\} \epsilon \Psi}{\text{g.l.b.}} p_0\left[x + \sum_{k=1}^{n} (y_k - x_k)\right], \tag{4.45}$$

$$p(x) \leqq Q_D(x) \equiv \underset{\{x_1, y_1\} \epsilon \Psi}{\text{g.l.b.}} p_0[x + (y_1 - x_1)] \tag{4.46}$$

is both necessary and sufficient; and the condition

$$p(x) \leqq R(x) \equiv \underset{n>0;\ \{x, y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[\frac{1}{n} \sum_{k=1}^{n} y_k\right] \tag{4.47}$$

is necessary.

It is essential to observe that Theorem 4.4 does not assert that $Q_A(x), \cdots, Q_D(x)$ are finite-valued. In many cases (for an example, see §5) one or more of $Q_A(x), \cdots, Q_D(x)$ is $-\infty$ for all $x \in E$; in such cases no $p \in P$ can exist satisfying (4.43), and it follows that no $f \in F$ exists satisfying (4.41) and (4.42). We observe also that if Ψ happens to have the property that $\{x, y\} \in \Psi$ implies $\{ax, ay\} \in \Psi$ for each $a \in R$, then $Q_A(x) = Q_B(x) = Q_C(x)$ for all $x \in E$.

Since $Q_A(x) \leqq Q_B(x) \leqq Q_D(x)$ and $Q_A(x) \leqq Q_C(x) \leqq Q_D(x)$ for each $x \in E$, we can prove necessity and sufficiency of (4.43), $\cdots$, (4.46) by proving necessity of (4.43) and sufficiency of (4.46). To prove necessity of (4.43), let $p \in P$ enforce (4.41) and (4.42), and let $x \in E$ be fixed. Let $n > 0$; $a_k \in R$; and let $\{x_k, y_k\} \in \Psi$. It follows from Theorem 4.3 that

$$p[\pm(y_k - x_k)] = 0 \qquad (k = 1, \cdots, n), \tag{4.481}$$

hence from Lemma 2.3 that

$$p\left[\pm \sum_{k=1}^{n} a_k(y_k - x_k)\right] = 0, \tag{4.482}$$

and therefore from Lemma 2.2 that

$$p(x) = p\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right]. \tag{4.483}$$

It follows from Theorem 4.1 that $p(\xi) \leqq p_0(\xi)$ for each $\xi \in E$; and if we let ξ be the argument of p in the right member of (4.483), we find

$$p(x) \leqq p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right]. \tag{4.484}$$

From (4.484) we obtain (4.43). To prove sufficiency of (4.46), let $p \in P$ satisfy (4.46). Putting $y_1 = x_1 = 0$, we see that $p(x) \leqq p_0(x)$ and it follows from Theorem 4.1 that p enforces (4.41). If $\{x, y\} \in \Psi$, then we can put $y_1 = y$, $x_1 = x$ in (4.46) to obtain

$$p(x - y) \leqq p_0[(x - y) + (y - x)] = p_0(0) = 0,$$

and put $y_1 = x$, $x_1 = y$ to obtain

$$p(y - x) \leqq p_0[(y - x) + (x - y)] = p_0(0) = 0.$$

It follows from Lemma 2.6 that $p[\pm(y - x)] = 0$ when $\{x, y\} \in \Psi$ and hence from Theorem 4.3 that p enforces (4.42). This completes the proof of necessity and sufficiency of (4.43), $\cdots$, (4.46). To prove necessity of (4.47) we observe that if, in the g.l.b. of (4.44), we require that $x_k = x$ for all k, the g.l.b. will not be

decreased and hence that $Q_B(x) \leqq R(x)$. Necessity of (4.47) follows, and proof of Theorem 4.4 is complete.

THEOREM 4.5. *If $p_0 \in P$ and*

$$Q_A(x) \equiv \underset{n>0;\, a_k \in R;\, \{x_k,y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right]$$

is finite for at least one $x_0 \in E$, then $Q_A(x)$ is a p-function (therefore finite-valued for all $x \in E$) which enforces the properties $f(x) \leqq p_0(x)$ and $f(y) = f(x)$ for $\{x, y\} \in \Psi$.

When x_0, $x \in E$; $a_k \in R$; and $\{x_k, y_k\} \in \Psi$, we have

$$Q_A(x_0) \leqq p_0\left[x_0 + \sum_{k=1}^{n} a_k(y_k - x_k)\right] = p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k) + (x_0 - x)\right]$$

$$\leqq p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right] + p_0(x_0 - x).$$

Hence

$$Q_A(x_0) - p_0(x_0 - x) \leqq p_0\left[x + \sum_{k=1}^{n} a_k(y_k - x_k)\right],$$

and we see that finiteness of $Q_A(x_0)$ and $p_0(x_0 - x)$ implies finiteness of $Q_A(x)$.

If $x \in E$ and $t > 0$, then

$$Q_A(tx) = \underset{n>0;\, a_k \in R;\, \{x_k,y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[tx + \sum_{k=1}^{n} a_k(y_k - x_k)\right]$$

$$= t \underset{n>0;\, a_k/t \in R;\, \{x_k,y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[x + \sum_{k=1}^{n} (a_k/t)(y_k - x_k)\right] = tQ_A(x).$$

If $x, y \in E$ and $\epsilon > 0$ are fixed, we can choose $m, a_1, \cdots, a_m \in R$ and $\{u_j, v_j\} \in \Psi$ such that

$$p_0\left[x + \sum_{j=1}^{m} a_j(v_j - u_j)\right] < Q_A(x) + \epsilon;$$

and choose $n, b_1, \cdots, b_n \in R$ and $\{x_k, y_k\} \in \Psi$ such that

$$p_0\left[y + \sum_{k=1}^{n} b_k(y_k - x_k)\right] < Q_A(y) + \epsilon.$$

It then follows from the definition of $Q_A(x + y)$ that

$$Q_A(x + y) \leqq p_0\left[(x + y) + \sum_{j=1}^{m} a_j(v_j - u_j) + \sum_{k=1}^{n} b_k(y_k - x_k)\right]$$

$$\leqq Q_A(x) + Q_A(y) + 2\epsilon.$$

Arbitrariness of $\epsilon > 0$ gives

$$Q_A(x + y) \leqq Q_A(x) + Q_A(y).$$

It now follows from Lemma 2.1 that $Q_A \in P$ and hence from Theorem 4.4 that Q_A enforces (4,41) and (4.42). This proves Theorem 4.5.

From Theorems 4.4 and 4.5 and Definition 4.13, we obtain

THEOREM 4.6. *If $Q_A(x) > -\infty$ for at least one $x \in E$, then $Q_A(x)$ is the greatest $p \in P$ which enforces* (4.41) *and* (4.42); *if $Q_A(x) = -\infty$ for at least one $x \in E$, then no $f \in F$ exists satisfying* (4.41) *and* (4.42).

THEOREM 4.7. *Let $p_0 \in P$ and Ψ be fixed. In order that $f \in F$ may exist with*

$$f(x) \leqq p_0(x); \quad f(y) = f(x) \qquad (x \in E;\ \{x, y\} \in \Psi), \tag{4.71}$$

it is necessary and sufficient that

$$Q_A(0) \equiv \underset{n>0;\, a_k \in R;\, \{x_k, y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[\sum_{k=1}^{n} a_k(y_k - x_k)\right] = 0. \tag{4.72}$$

To prove necessity, suppose $f \in F$ exists satisfying (4.71). Then by Theorem 1.3, $p \in P$ exists and this enforces (4.71). Then, by Theorem 4.4, $p(x) \leqq Q_A(x)$ so $Q_A(x)$ must be finite-valued for all $x \in E$. Hence Theorem 4.5 implies $Q_A \in P$ and therefore $Q_A(0) = 0$. This proves necessity of (4.72). Sufficiency of (4.72) follows from Theorem 4.5.

THEOREM 4.8. *Let $p_0 \in P$ and Ψ be fixed. If*

$$Q_D(x) = \underset{\{x_1, y_1\} \in \Psi}{\text{g.l.b.}} p_0[x + (y_1 - x_1)]$$

is finite for at least one $x_0 \in E$, then it is finite for all $x \in E$.

Proof of Theorem 4.8 is analogous to the first part of our proof of Theorem 4.5 and is left to the reader. It is easy to show that $Q_D(tx) = tQ_D(x)$ for $t > 0$, $x \in E$ in case $\{x, y\} \in \Psi$ implies $\{ax, ay\} \in \Psi$ for each $a \in R$; and to show that $Q_D(x + y) \leqq Q_D(x) + Q_D(y)$ in case $\{x_1, y_1\} \in \Psi$ and $\{x_2, y_2\} \in \Psi$ implies $\{x_1 + x_2, y_1 + y_2\} \in \Psi$. However, if Ψ has these two properties, then $Q_D(x) = Q_A(x)$ for all $x \in E$.

5. **An illustrative example.** In this section, let E denote the linear space of real bounded functions $x \equiv x(s)$ defined over $-\infty < s < \infty$, and let

$$p_0(x) = \overline{\lim_{s \to \infty}}\, x(s) \qquad (x \in E). \tag{5.01}$$

Observe that, as the notation implies, $p_0 \in P$. Let

$$\varphi(s) = s + (1/12) \sin 2\pi s \qquad (-\infty < s < \infty). \tag{5.02}$$

The function $\varphi(s)$ is analytic and has a positive derivative which is bounded and bounded from zero. In fact, $1 - \pi/6 \leqq \varphi'(s) \leqq 1 + \pi/6$. Finally, let Ψ consist of all pairs $\{x(s), x(s + \lambda)\}$ with $x \in E$, $\lambda \in R$ together with all pairs $\{x(s), x(\varphi(s))\}$ and $\{x(\varphi(s)), x(s)\}$ with $x \in E$. Observe that, for this example, $Q_A(x) = Q_B(x) = Q_C(x)$ for all $x \in E$. We proceed to show that, for this example,

$$Q_A(0) = \underset{n>0;\, \{x_k, y_k\} \in \Psi}{\text{g.l.b.}} p_0\left[\sum_{k=1}^{n} (y_k - x_k)\right] = -\infty. \tag{5.1}$$

If $M \epsilon R$, $[s]$ denotes the greatest integer $\leqq s$, and $\xi(s) \epsilon E$ is defined by

$$(5.21)\qquad \begin{aligned} \xi(s) &= M && (s-[s] < \tfrac{1}{3}),\\ &= 0 && (\text{otherwise}), \end{aligned}$$

then, since $\varphi(s)$ is increasing, $\varphi(s+1) = \varphi(s)$; $\varphi(0) = 0$, and $\varphi(\frac{1}{4}) = \frac{1}{3}$,

$$(5.22)\qquad \begin{aligned} \xi(\varphi(s)) &= M && (s-[s] < \tfrac{1}{4}),\\ &= 0 && (\text{otherwise}). \end{aligned}$$

Therefore

$$(5.23)\qquad \begin{aligned} \xi(\varphi(s)) - \xi(s) &= -M && (\tfrac{1}{4} \leqq s-[s] < \tfrac{1}{3}),\\ &= 0 && (\text{otherwise}), \end{aligned}$$

and hence

$$(5.24)\qquad \sum_{k=1}^{12}\left[\xi\left(\varphi\left(s + \frac{k}{12}\right)\right) - \xi\left(s + \frac{k}{12}\right)\right] = -M \quad (-\infty < s < \infty),$$

or

$$\sum_{k=1}^{12}\left[\left\{\xi\left(\varphi\left(s + \frac{k}{12}\right)\right) - \xi(\varphi(s))\right\} + \{\xi(\varphi(s)) - \xi(s)\} + \left\{\xi(s) - \xi\left(s + \frac{k}{12}\right)\right\}\right] = -M.$$

The last equality has the form

$$(5.25)\qquad \sum_{k=1}^{n}[y_k(s) - x_k(s)] = -M \qquad (-\infty < s < \infty),$$

where $\{x_k, y_k\} \epsilon \Psi$. The definition (5.01) of p_0 now gives

$$(5.26)\qquad p_0 \sum_{k=1}^{n}[x_k(s) - y_k(s)] = -M.$$

It follow from (5.26) that the left member of (5.1) is $\leqq -M$ and, since $M \epsilon R$ is arbitrary, that (5.1) holds.

Since (5.1) holds, there is no $p \epsilon P$ satisfying (4.43) and hence by Theorem 4.4 there is no $f \epsilon F$ with the properties

$$(5.31)\qquad f(x) \leqq \overline{\lim_{s\to\infty}}\, x(s) \qquad (x \epsilon E),$$

$$(5.32)\qquad f(x(s+\lambda)) = f(x(s)) \qquad (\lambda \epsilon R, x \epsilon E),$$

$$(5.33)\qquad f(x(\varphi(s))) = f(x(s)) \qquad (x \epsilon E).$$

It follows that there exists no generalized limit "$\operatorname*{Lim}_{s\to\infty} x(s)$", defined for all $x \epsilon E$, having the properties

$$(5.41)\qquad \operatorname*{Lim}_{s\to\infty}[ax(s) + by(s)] = a \operatorname*{Lim}_{s\to\infty} x(s) + b \operatorname*{Lim}_{s\to\infty} y(s) \qquad (a, b \epsilon R; x, y \epsilon E),$$

$$(5.42)\qquad \underline{\lim_{s\to\infty}}\, x(s) \leqq \operatorname*{Lim}_{s\to\infty} x(s) \leqq \overline{\lim_{s\to\infty}}\, x(s) \qquad (x \epsilon E),$$

$$\operatorname*{Lim}_{s\to\infty} x(s + \lambda) = \operatorname*{Lim}_{s\to\infty} x(s) \qquad (\lambda \in R;\ x \in E), \tag{5.43}$$

$$\operatorname*{Lim}_{s\to\infty} x(\varphi(s)) = \operatorname*{Lim}_{s\to\infty} x(s) \qquad (x \in E). \tag{5.44}$$

For if such a "Lim" exists we could put $f(x) = \operatorname*{Lim}_{s\to\infty} x(s)$ to obtain $f \in F$ satisfying (5.31), (5.32), and (5.33).

For the example under consideration we have, when $\{x, y\} \in \Psi$, either

$$p_0(y - x) \equiv \overline{\lim_{s\to\infty}}\ [\pm x(s + \lambda) \mp x(s)] \geqq 0,$$

or

$$p_0(y - x) \equiv \overline{\lim_{s\to\infty}}\ [\pm x(\varphi(s)) \mp x(s)] \geqq 0.$$

These inequalities and the equality $p_0(0) = 0$ give

$$Q_D(0) = \operatorname*{g.l.b.}_{\{x,y\} \in \Psi} p_0(y - x) = 0,$$

and it follows from Theorem 4.8 that $Q_D(x)$ is finite for all $x \in E$.

From (5.1) and Theorem 4.4 it follows that there is no $p \in P$ satisfying any one of the four conditions (4.43), $\cdots$, (4.46). This example therefore shows that finiteness of $Q_D(x)$ does not guarantee existence of $f \in F$ with $f(x) \leqq Q_D(x)$, i.e., does not guarantee that $Q_D(x)$ is an r-function. In particular, Theorems 4.5, 4.6 and 4.7 would not hold if $Q_A(x)$ were replaced by $Q_D(x)$ in their statements.

6. **Functionals $f \in F$ invariant under** G. Let G represent a group of transformations $g \equiv g(x)$ each of which maps E univalently into itself and is linear, i.e.,

$$g(ax + by) = ag(x) + bg(y) \qquad (g \in G;\ a, b \in R;\ x, y \in E). \tag{6.01}$$

We shall at times omit parentheses, writing g_1g_2x for $g_1(g_2(x))$, pgx for $p(g(x))$, etc.

Let Ψ be the set of all pairs $\{x, g(x)\}$ with $g \in G$, $x \in E$. Linearity of g implies that $\{ax, ag(x)\} \in \Psi$ when $a \in R$, $g \in G$, $x \in E$. Thus Ψ has the property that if $\{x, y\} \in \Psi$, then $\{ax, ay\} \in \Psi$ for each $a \in R$. Therefore, under the definitions (4.43), (4.44), and (4.45), we have

$$Q_A(x) = Q_B(x) = Q_C(x) \qquad (x \in E). \tag{6.02}$$

These considerations, together with Theorems 4.4, 4.5 and 4.6 give the following theorem:

THEOREM 6.1. *Let $p_0 \in P$ and G be fixed. In order that $p \in P$ may enforce the properties*

$$f(x) \leqq p_0(x) \qquad (x \in E), \tag{6.11}$$

$$f(g(x)) = f(x) \qquad (g \in G;\ x \in E), \tag{6.12}$$

each of the two conditions

$$(6.13)\qquad p(x) \leqq Q_B(x) \equiv \underset{n>0;\, g_k \in G;\, x_k \in E}{\text{g.l.b.}} p_0\left[x + \frac{1}{n}\sum_{k=1}^{n}(g_k x_k - x_k)\right] \qquad (x \in E),$$

$$(6.14)\qquad p(x) \leqq Q_D(x) \equiv \underset{g_1 \in G;\, x_1 \in E}{\text{g.l.b.}} p_0[x + (g_1 x_1 - x_1)] \qquad (x \in E)$$

is both necessary and sufficient, and the condition

$$(6.15)\qquad p(x) \leqq R(x) \equiv \underset{n>0;\, g_k \in G}{\text{g.l.b.}} p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \qquad (x \in E)$$

is necessary. If $Q_B(x) = -\infty$ for at least one $x \in E$, then no $f \in F$ exists satisfying (6.11) *and* (6.12). *If $Q_B(x) > -\infty$ for at least one $x \in E$, then $Q_B(x)$ is the greatest p-function which enforces* (6.11) *and* (6.12).

We digress to state and prove a lemma which we shall use in §8.

LEMMA 6.2. *If Γ is a subgroup of G; $q_1(x)$ is a functional defined over E; and $q_2(x)$, $q_3(x)$ and $q_4(x)$ are defined (finite or $-\infty$) over E by the formulas*

$$(6.21)\qquad q_2(x) = \underset{m>0;\, \gamma_j \in \Gamma}{\text{g.l.b.}} q_1\left[\frac{1}{m}\sum_{j=1}^{m} \gamma_j x\right],$$

$$(6.22)\qquad q_3(x) = \underset{n>0;\, g_k \in G}{\text{g.l.b.}} q_2\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right],$$

$$(6.23)\qquad q_4(x) = \underset{n>0;\, g_k \in G}{\text{g.l.b.}} q_1\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right],$$

then $q_3(x) = q_4(x)$ for all $x \in E$.

To prove this lemma, we observe first that setting $m = 1$ and $\gamma_1 = I$ (the identity of G) in the argument of q_1 in (6.21) shows that $q_2(x) \leqq q_1(x)$; it then follows on comparing (6.22) and (6.23) that $q_3(x) \leqq q_4(x)$.

On the other hand it follows from (6.21) and the hypothesis $\Gamma \subset G$ that

$$q_2(\xi) \geqq \underset{m>0;\, \gamma_j \in G}{\text{g.l.b.}} q_1\left[\frac{1}{m}\sum_{j=1}^{m} \gamma_j \xi\right] \qquad (\xi \in E)$$

and hence that

$$q_2\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \geqq \underset{m>0;\, \gamma_j \in G}{\text{g.l.b.}} q_1\left[\frac{1}{m}\sum_{j=1}^{m} \gamma_j \frac{1}{n}\sum_{k=1}^{n} g_k x\right].$$

Taking g.l.b. for $n > 0$, $g_k \in G$, and using (6.22), linearity of the transformations in G, and finally (6.23), we obtain the estimate,

$$\begin{aligned} q_3(x) &\geqq \underset{n>0;\, g_k \in G}{\text{g.l.b.}}\ \underset{m>0;\, \gamma_j \in G}{\text{g.l.b.}} q_1\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} \gamma_j g_k x\right] \\ &= \underset{m,n>0;\, g_k, \gamma_j \in G}{\text{g.l.b.}} q_1\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} \gamma_j g_k x\right] \geqq q_4(x). \end{aligned}$$

Combining inequalities gives $q_3(x) = q_4(x)$ and Lemma 6.2 is proved.

7. **Both $p_0 \in P$ and $f \in F$ invariant under G.** Theorem 6.1 does not guarantee that $Q_D(x)$ and $R(x)$, as defined by (6.14) and (6.15), are not $-\infty$; however, if $p_0(x)$ is invariant under G, then they must be finite-valued in accordance with

THEOREM 7.1. *If $p_0 \in P$ and G are so related that*

$$p_0(g(x)) = p_0(x) \qquad (g \in G,\ x \in E), \tag{7.11}$$

then

$$Q_D(x) \equiv \underset{g_1 \in G;\, x_1 \in E}{\text{g.l.b.}}\ p_0[x + (g_1x_1 - x_1)], \tag{7.12}$$

$$R(x) \equiv \underset{n>0;\, g_k \in G}{\text{g.l.b.}}\ p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \tag{7.13}$$

are finite-valued for all $x \in E$ and

$$-p_0(-x) \leqq Q_D(x) \leqq p_0(x) \qquad (x \in E), \tag{7.14}$$

$$-p_0(-x) \leqq R(x) \leqq p_0(x) \qquad (x \in E). \tag{7.15}$$

Setting $x_1 = 0$ in the argument of p_0 in (7.12) shows that $Q_D(x) \leqq p_0(x)$. On the other hand, if $g_1 \in G$, $x_1 \in E$, then (7.11) and Lemma 2.41 give

$$0 \leqq p_0[g_1x_1 - x_1].$$

Hence

$$\begin{aligned} 0 &\leqq p_0[(x + g_1x_1 - x_1) - x] \\ &\leqq p_0[x+(g_1x_1 - x_1)] + p_0(-x), \end{aligned}$$

and $-p_0(-x) \leqq Q_D(x)$ follows. This proves (7.14).

Putting $n = 1$ and $g_1 = I$, the identity of G, in the argument of p_0 in (7.13) shows that $R(x) \leqq p_0(x)$. On the other hand, (7.13) implies

$$-R(x) = \underset{n>0;\, g_k \in G}{\text{l.u.b.}}\ -p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right].$$

Using Lemma 2.5, linearity of $g \in G$, and (7.11), we continue to obtain

$$\begin{aligned} -R(x) &\leqq \underset{n>0;\, g_k \in G}{\text{l.u.b.}}\ p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k(-x)\right] \\ &\leqq \underset{n>0;\, g_k \in G}{\text{l.u.b.}}\ \frac{1}{n}\sum_{k=1}^{n} p_0 g_k(-x) = p_0(-x), \end{aligned}$$

so that $-p_0(-x) \leqq R(x)$. This completes the proof of (7.15) and hence of Theorem 7.1.

In case p_0 is invariant under G, we can strengthen Theorem 6.1 by proving that (6.15) is sufficient to ensure that p enforces (6.11) and (6.12). This fact and Theorem 6.1 give

THEOREM 7.2. *Let $p_0 \in P$ and G be so related that*

$$p_0(g(x)) = p_0(x) \qquad (g \in G;\ x \in E). \tag{7.21}$$

In order that $p \in P$ may enforce the properties

$$f(x) \leqq p_0(x) \qquad (x \in E), \tag{7.22}$$

$$f(g(x)) = f(x) \qquad (g \in G;\, x \in E), \tag{7.23}$$

each of the three conditions

$$p(x) \leqq Q_B(x) \equiv \underset{n>0;\, g_k \in G;\, x_k \in E}{\text{g.l.b.}} p_0\left[x + \frac{1}{n}\sum_{k=1}^{n}(g_k x_k - x_k)\right] \qquad (x \in E), \tag{7.24}$$

$$p(x) \leqq Q_D(x) \equiv \underset{g_1 \in G,\, x_1 \in E}{\text{g.l.b.}} p_0[x + (g_1 x_1 - x_1)] \qquad (x \in E), \tag{7.25}$$

$$p(x) \leqq R(x) \equiv \underset{n>0;\, g_k \in G}{\text{g.l.b.}} p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \qquad (x \in E), \tag{7.26}$$

is both necessary and sufficient.

To prove sufficiency of (7.26), let $p \in P$ satisfy (7.26). Then (7.26) and Theorem 7.1 give $p(x) \leqq R(x) \leqq p_0(x)$ so that, by Theorem 4.1, p enforces (7.22). To prove that p enforces (7.23), let $g \in G$ and $x \in E$ be fixed. Then (7.26) gives

$$p(gx - x) \leqq \underset{n>0;\, g_k \in G}{\text{g.l.b.}} p_0\left[\frac{1}{n}\sum_{k=1}^{n}(g_k gx - g_k x)\right].$$

The g.l.b. on the right will not be decreased if we require that $g_k = g^{k-1}$ (g^0 being the identity of G) for each k. Hence

$$\begin{aligned} p(gx - x) &\leqq \underset{n>0}{\text{g.l.b.}}\, p_0\left[\frac{1}{n}\sum_{k=1}^{n}(g^k x - g^{k-1}x)\right] \\ &= \underset{n>0}{\text{g.l.b.}}\, \frac{1}{n} p_0[g^n x - x] \leqq \underset{n>0}{\text{g.l.b.}}\, \frac{1}{n}[p_0(g^n x) + p_0(-x)] \\ &= \underset{n>0}{\text{g.l.b.}}\, \frac{1}{n}[p_0(x) + p_0(-x)] = 0. \end{aligned} \tag{7.27}$$

Since a similar argument shows that $p(x - gx) \leqq 0$, it follows from Lemma 2.6 that $p[\pm(gx - x)] = 0$ and hence from Theorem 4.3 that p enforces (7.23). This completes the proof of Theorem 7.2.

The condition (7.21) is not sufficient to ensure that $Q_B(x)$ as defined by (7.24) is finite-valued. To see this, let E, p_0, φ, and Ψ be specialized as in the example of §5. Let G be the group generated by transformations of the form $gx(s) = x(s + \lambda)$ with $\lambda \in R$ and the transformation $gx(s) = x(\varphi(s))$. Using results of §5, it is easy to see that for this example $Q_B(x)$ defined by (7.24) is $-\infty$ for all $x \in E$. It thus appears that the $Q_B(x)$ of Theorem 7.2, which by Theorem 6.1 must be a p-function if it is finite-valued, need not be finite-valued. In such cases no $f \in F$ satisfying (7.22) and (7.23) exists; and no $p \in P$ satisfying (7.24) or (7.25) or (7.26) exists. Thus we see that $Q_D(x)$ and $R(x)$, which must be finite-valued by Theorem 7.1, may fail to be r-functions and hence fail to be p-functions even when (7.21) holds.

On the other hand it follows from our theorems that if $Q_B(x)$ is finite-valued, or if $Q_D(x)$ is an r-function, or if $R(x)$ is an r-function, then all must be true and $f \in F$ exists satisfying (7.22) and (7.23).

It is obvious from the definitions (7.24), (7.25), and (7.26) that $Q_B(x) \leqq Q_D(x)$ and $Q_B(x) \leqq R(x)$ for each $x \in E$. We do not stop here to consider the interesting possibility that condition (7.21) may imply either or both of the inequalities $Q_D(x) \leqq R(x)$ and $R(x) \leqq Q_D(x)$ for all $x \in E$.

The following theorem enables us to add two criteria, namely, $p(x) \leqq Q_B^{(1)}(x)$ and $p(x) \leqq Q_B^{(2)}(x)$, to the set (7.24), (7.25), (7.26) in Theorem 7.2.

THEOREM 7.3. *If $p_0 \in P$ and G are so related that*

$$p_0(g(x)) = p_0(x) \qquad (g \in G;\ x \in E), \tag{7.31}$$

and $Q_B(x)$, $Q_B^{(1)}(x)$ and $Q_B^{(2)}(x)$ are defined for $x \in E$ by

$$Q_B(x) = \underset{n>0;\, g_k \in G;\, x_k \in E}{\text{g.l.b.}}\ p_0\left[x + \frac{1}{n}\sum_{k=1}^{n}(g_k x_k - x_k)\right], \tag{7.32}$$

$$Q_B^{(1)}(x) = \underset{m>0;\, \Sigma x_j = x}{\text{g.l.b.}}\ \sum_{j=1}^{m}\ \underset{g_j \in G;\, \xi_j \in E}{\text{g.l.b.}}\ p_0[x_j + (g_j\xi_j - \xi_j)], \tag{7.33}$$

$$Q_B^{(2)}(x) = \underset{m>0;\, \Sigma x_j = x}{\text{g.l.b.}}\ \sum_{j=1}^{m}\ \underset{n_j>0;\, g_{jk} \in G}{\text{g.l.b.}}\ p_0\left[\frac{1}{n_j}\sum_{k=1}^{n_j} g_{jk} x_j\right], \tag{7.34}$$

then, whether the bounds be finite or $-\infty$,

$$Q_B(x) = Q_B^{(1)}(x) = Q_B^{(2)}(x) \qquad (x \in E). \tag{7.35}$$

To prove Theorem 7.3, we observe first that, with $Q_D(x)$ and $R(x)$ defined by (7.25) and (7.26),

$$Q_B^{(1)}(x) = \underset{m>0;\, \Sigma x_j = x}{\text{g.l.b.}}\ \sum_{j=1}^{m} Q_D(x_j), \tag{7.36}$$

$$Q_B^{(2)}(x) = \underset{m>0;\, \Sigma x_j = x}{\text{g.l.b.}}\ \sum_{j=1}^{m} R(x_j). \tag{7.37}$$

Since

$$Q_D(tx) = tQ_D(x); \quad R(tx) = tR(x), \qquad (t > 0;\ x \in E)' \tag{7.38}$$

it follows from Theorem 3.3 that (7.25) holds if and only if $p(x) \leqq Q_B^{(1)}(x)$, and that (7.26) holds if and only if $p(x) \leqq Q_B^{(2)}(x)$. Consider now two cases. In case $f \in F$ exists satisfying (7.22) and (7.23), then by Theorems 7.2 and 6.1, $Q_B(x)$ is a p-function while $Q_D(x)$ and $R(x)$ are r-functions. It follows from Theorems 7.2 and 3.2 that $Q_B(x)$, $Q_B^{(1)}(x)$, and $Q_B^{(2)}(x)$ each represent the greatest p-function which enforces (7.22) and (7.23). This gives (7.35) for the first case. In case no $f \in F$ exists satisfying (7.22) and (7.23), it follows from Theorems 7.2, 6.1, and 3.4 that each member of (7.35) is $-\infty$ for all $x \in E$. This completes the proof of Theorem 7.3.

The proof we have just given illustrates the fact that the concept of "greatest $p \epsilon P$ which enforces $\cdots$" can lead to discovery of equality of functionals which might appear to be unequal (or which might possibly be so complicated that nothing would be apparent).

8. **Both $p_0 \epsilon P$ and $f \epsilon F$ invariant under solvable G.** In §7 we saw that $R(x)$, defined by (7.13), may fail to be an r-function and hence may fail to be a p-function even when (7.11) holds. The fundamental result of §8 is that if G is solvable and (7.11) holds, then $R(x)$ must be a p-function. We denote the derived subgroup of G (the subgroup of G generated by the commutators $g_1g_2g_1^{-1}g_2^{-1}$) by G', the derived group of G' by G'', etc. A group is called solvable if some derived group $G^{(r)}$ consists of the identity alone.

THEOREM 8.1. *If G is solvable and $p_0 \epsilon P$ has the property*

$$p_0gx = p_0x \qquad (g \epsilon G,\ x \epsilon E), \tag{8.11}$$

then the functional $R(x)$ defined by

$$R(x) = \underset{n>0;\, g_k \epsilon G}{\text{g.l.b.}}\ p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \qquad (x \epsilon E) \tag{8.12}$$

is a p-function.

To prove this theorem, let $G_0 \subset G_1 \subset G_2 \subset \cdots \subset G_r$ be a sequence of groups having the properties:

1. G_0 consists of the identity alone;
2. for each $\alpha = 1, 2, \cdots, r$, $G_{\alpha-1}$ is the derived group G'_α of G_α; and
3. G_r is the group G appearing in Theorem 8.1.

With p_0 given in Theorem 8.1, we define $p_1(x), p_2(x), \cdots, p_r(x)$ by the recursion formulas

$$p_\alpha(x) = \underset{n>0;\, g_k \epsilon G_\alpha}{\text{g.l.b.}}\ p_{\alpha-1}\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \qquad (\alpha = 1, 2, \cdots, r). \tag{8.13}$$

It follows from Theorem 7.1 that $p_\alpha(x)$ is finite-valued for each $\alpha = 1, 2, \cdots, r$ and $x \epsilon E$. Our proof of Theorem 8.1 depends upon two lemmas.

LEMMA 8.2. *For each $\alpha = 1, 2, \cdots, r$*

$$p_\alpha(x) = \underset{n>0;\, g_k \epsilon G_\alpha}{\text{g.l.b.}}\ p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \qquad (x \epsilon E). \tag{8.14}$$

The above lemma follows from the definitions (8.13) and iterated use of Lemma 6.2.

LEMMA 8.3. *For each $\alpha = 0, 1, 2, \cdots, r$, $p_\alpha(x)$ is a p-function with the property*

$$p_\alpha[\pm(gx - x)] = 0 \qquad (g \epsilon G_\alpha\,;\ x \epsilon E), \tag{8.31}$$

and if $\alpha < r$, then p_α has the property

$$p_\alpha gx = p_\alpha x \qquad (g \epsilon G_{\alpha+1};\ x \epsilon E). \tag{8.32}$$

Our hypotheses give $p_0 \in P$ and imply that (8.31) and (8.32) hold when $\alpha = 0$. To complete an induction proof of Lemma 8.3, let $0 < \alpha \leqq r$, $p_{\alpha-1} \in P$, and

$$p_{\alpha-1}[\pm(gx - x)] = 0 \qquad (g \in G_{\alpha-1}\,;\, x \in E), \tag{8.331}$$

$$p_{\alpha-1}gx = p_{\alpha-1}x \qquad (g \in G_{\alpha}\,;\, x \in E). \tag{8.332}$$

We have seen that p_α is finite-valued, i.e., is a functional. We show next that $p_\alpha \in P$.

It follows easily from (8.13) that $p_\alpha(tx) = tp_\alpha(x)$ when $t > 0$ and $x \in E$. To prove $p_\alpha(x + y) \leqq p_\alpha x + p_\alpha y$, let $x, y \in E$ and $\epsilon > 0$ be fixed. Choose $g_1, \cdots, g_m, h_1, \cdots, h_n \in G_\alpha$ such that

$$p_{\alpha-1}\left[\frac{1}{m}\sum_{j=1}^{m} g_j x\right] < p_\alpha(x) + \epsilon, \tag{8.341}$$

$$p_{\alpha-1}\left[\frac{1}{n}\sum_{k=1}^{n} h_k y\right] < p_\alpha(y) + \epsilon. \tag{8.342}$$

Using linearity of the elements of G, the hypothesis $p_{\alpha-1} \in P$, and (8.332), we obtain the estimate

$$p_{\alpha-1}\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} h_k g_j x\right] = p_{\alpha-1}\left[\frac{1}{n}\sum_{k=1}^{n} h_k \frac{1}{m}\sum_{j=1}^{m} g_j x\right]$$

$$\leqq \frac{1}{n}\sum_{k=1}^{n} p_{\alpha-1}\left[h_k \frac{1}{m}\sum_{j=1}^{m} g_j x\right] = \frac{1}{n}\sum_{k=1}^{n} p_{\alpha-1}\left[\frac{1}{m}\sum_{j=1}^{m} g_j x\right].$$

But the last member of this equality is the left member of (8.341); hence

$$p_{\alpha-1}\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} h_k g_j x\right] < p_\alpha(x) + \epsilon. \tag{8.351}$$

An analogous estimate together with (8.342) gives

$$p_{\alpha-1}\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} g_j h_k y\right] < p_\alpha(y) + \epsilon. \tag{8.352}$$

Our hypothesis (8.331) implies that

$$p_{\alpha-1}[\pm(hgh^{-1}g^{-1}\xi - \xi)] = 0 \qquad (h, g \in G_\alpha;\, \xi \in E),$$

and hence, since the elements of G map E univalently into itself, that

$$p_{\alpha-1}[\pm(hgx - ghx)] = 0 \qquad (h, g \in G_\alpha;\, x \in E).$$

It follows from g_j, $h_k \in G_\alpha$ and Lemma 2.3 that

$$p_{\alpha-1}\left[\pm\sum_{j,k=1}^{m,n} (h_k g_j x - g_j h_k x)\right] = 0.$$

This equality, Lemma 2.4, and (8.351) imply

$$p_{\alpha-1}\left[\frac{1}{mn}\sum_{j,k=1}^{m,n} g_j h_k x\right] < p_\alpha(x) + \epsilon. \tag{8.353}$$

Using the definition (8.13) and formulas (8.352) and (8.353), we find

$$p_\alpha(x + y) \leqq p_{\alpha-1}\left[\frac{1}{mn} \sum_{j,k=1}^{m,n} g_j h_k(x + y)\right] \leqq p_\alpha x + p_\alpha y + 2\epsilon.$$

Arbitrariness of $\epsilon > 0$ gives $p_\alpha(x + y) \leqq p_\alpha x + p_\alpha y$. This proves $p_\alpha \in P$.

If $g \in G_\alpha$, it follows from the definition (8.13) that

$$p_\alpha(gx - x) \leqq \underset{n>0}{\text{g.l.b.}}\ p_{\alpha-1}\left[\frac{1}{n} \sum_{k=1}^{n} (g^k x - g^{k-1} x)\right],$$

and, using (8.332), we can proceed as in (7.27) to obtain $p_\alpha(gx - x) \leqq 0$. Likewise $p_\alpha(x - gx) \leqq 0$ and (8.31) follows.

We next prove (8.32). Let $g \in G_{\alpha+1}$ and $x \in E$ be fixed. It follows from Lemma 8.2 that

$$p_\alpha(\xi) = \underset{n>0;\, g_k \in G_\alpha}{\text{g.l.b.}}\ p_0\left[\frac{1}{n} \sum_{k=1}^{n} g_k \xi\right] \qquad (\xi \in E). \tag{8.36}$$

Replacing ξ by gx in (8.36), and using (8.11) with g replaced by g^{-1}, we find

$$p_\alpha gx = \underset{n>0;\, g_k \in G_\alpha}{\text{g.l.b.}}\ p_0\left[\frac{1}{n} \sum_{k=1}^{n} g^{-1} g_k g x\right]. \tag{8.37}$$

Since $g \in G_{\alpha+1}$, we conclude that $g^{-1} g_k g \equiv (g^{-1} g_k g g_k^{-1}) g_k \in G_\alpha$ when $g_k \in G_\alpha$; hence

$$p_\alpha gx \geqq \underset{n>0;\, \gamma_k \in G_\alpha}{\text{g.l.b.}}\ p_0\left[\frac{1}{n} \sum_{k=1}^{n} \gamma_k x\right]. \tag{8.38}$$

Thus

$$p_\alpha gx \geqq p_\alpha x \qquad (g \in G_{\alpha+1};\, x \in E). \tag{8.39}$$

If in (8.39) we replace g by g^{-1} and x by gx, we obtain

$$p_\alpha x \geqq p_\alpha gx \qquad (g \in G_{\alpha+1};\, x \in E). \tag{8.391}$$

From (8.39) and (8.391) we obtain (8.32) and proof of Lemma 8.3 is complete.

Proof of Theorem 8.1 now follows easily. Definition (8.12) and Lemma 8.2 imply that $R(x) = p_\tau(x)$ and Lemma 8.3 therefore gives $R(x) \equiv p_\tau(x) \in P$. This proves Theorem 8.1.

Combining Theorems 7.2 and 8.1 we obtain

THEOREM 8.4. *If G is solvable and $p_0 \in P$ has the property*

$$p_0 gx = p_0 x \qquad (g \in G;\, x \in E), \tag{8.41}$$

then there exists $f \in F$ with the properties

$$f(x) \leqq p_0(x); \qquad f(g(x)) = f(x), \qquad (g \in G;\, x \in E), \tag{8.42}$$

and the functional

$$R(x) = \underset{n>0;\, g_k \in G}{\text{g.l.b.}}\ p_0\left[\frac{1}{n} \sum_{k=1}^{n} g_k x\right] \tag{8.43}$$

is the greatest p-function which enforces (8.42).

If G is solvable and (7.21) holds, then, by Theorem 8.1, $R(x)$ is a p-function; it follows from Theorem 7.2 that $Q_B(x)$, defined by (7.24), must be finite-valued and hence must be a p-function; and that $Q_D(x)$ must be an r-function. These observations and Theorem 8.4 enable us to strengthen the conclusions of Theorem 7.3 when G is solvable. Leaving details of proof to the reader, we give

THEOREM 8.5. *If G is solvable and $p_0 \in P$ has the property*

$$p_0 g x = p_0 x \qquad (g \in G;\ x \in E), \tag{8.51}$$

then the three functionals defined for $x \in E$ by

$$Q_B(x) = \operatorname*{g.l.b.}_{n>0;\ g_k \in G;\ x_k \in E} p_0\left[x + \frac{1}{n}\sum_{k=1}^{n}(g_k x_k - x_k)\right], \tag{8.52}$$

$$Q_B^{(1)}(x) = \operatorname*{g.l.b.}_{m>0;\ \Sigma x_j = x} \sum_{j=1}^{m} \operatorname*{g.l.b.}_{g_j \in G;\ \xi_j \in E} p_0[x_j + g_j\xi_j - \xi_j], \tag{8.53}$$

$$R(x) = \operatorname*{g.l.b.}_{n>0;\ g_k \in G} p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k x\right] \tag{8.54}$$

are p-functions (and therefore finite-valued) which are equal for all $x \in E$.

Theorem 8.5 is a theorem of analysis which may have some interest apart from its connection with the theory of linear functionals. We illustrate this remark by a very special example. Let E be the linear space of real bounded functions $x \equiv x(s)$; let $p_0(x) = \overline{\lim}_{s\to\infty} x(s)$; and let G be the group of transformations of the form $gx(s) = x(\mu s + \lambda)$ where $\mu, \lambda \in R$ with $\mu > 0$. It is easy to verify that the hypotheses of Theorem 8.5 are satisfied and it follows that $Q_B(0) \geqq 0$; in fact $Q_B(0) = 0$. It is also easy to verify that the following theorem is merely a statement that $Q(0) \geqq 0$ when E, p_0 and G are as we have specialized them.

THEOREM 8.6. *If $n > 0$; $x_1(s), \cdots, x_n(s)$ are real bounded functions; $\mu_1, \cdots, \mu_n > 0$; and $\lambda_1, \cdots, \lambda_n$ are real, then*

$$\overline{\lim_{s\to\infty}} \sum_{k=1}^{n} [x_k(\mu_k s + \lambda_k) - x_k(s)] \geqq 0. \tag{8.61}$$

This incidental corollary of Theorem 8.5 may be known; it is given here merely as a representative of a set of corollaries of Theorem 8.5 which seem interesting and challenge one to find a simple direct proof.

We point out finally in connection with the theorems of §8 involving solvable groups that the conclusions of Theorems 8.1, 8.4 and 8.5 may hold (at least for some $p_0 \in P$) when G is not solvable. In case $p_0 gx = p_0 x$ and p_0 happens to be linear, the right members of (8.12), (8.43), (8.52), (8.53) and (8.54) all reduce to the p-function p_0, irrespective of the character of G. In this trivial case, $f \in F$ satisfying (8.42) is obtained merely by setting $f = p_0$. It thus appears that a necessary and sufficient condition that (8.11) imply that $R(x)$ be a p-function cannot be given in terms of p_0 alone or in terms of G alone, but must involve a correlation of p_0 and G. In view of this fact, the following theorem may be of some interest.

THEOREM 8.7. *If G (solvable or not) and $p_0 \in P$ are so related that $R(x) \in P$, then for each linear transformation γ (not necessarily $\in G$) with domain and range in E the functional $R_\gamma(x)$ defined by*

$$R_\gamma(x) = \underset{n>0;\, g_k \in G}{\text{g.l.b.}}\; p_0\left[\frac{1}{n}\sum_{k=1}^{n} g_k \gamma x\right] \qquad (x \in E) \tag{8.71}$$

is a p-function.

It follows from (8.43) and (8.71) that $R_\gamma(x) = R(\gamma x)$; hence an application of Lemma 2.9 gives Theorem 8.7. In case γ has an inverse and $p_0\gamma^{-1}x = p_0 x$ for each $x \in E$, (8.71) can be written

$$R_\gamma(x) = \underset{n>0;\, g_k \in G}{\text{g.l.b.}}\; p_0\left[\frac{1}{n}\sum_{k=1}^{n} \gamma^{-1} g_k \gamma x\right] \qquad (x \in E). \tag{8.72}$$

Observe that as g_k ranges over G, $\gamma^{-1}g_k\gamma$ ranges over a group Γ which is simply isomorphic with G and which is therefore solvable if and only if G is solvable.

9. **Applications to limits.** Let E denote the set of real bounded functions $x = x(s)$ defined over $-\infty < s < \infty$. As in the example of §1, we write $\underset{s\to\infty}{\text{Lim}}\, x(s)$ in place of a linear functional $f(x)$. Thus we have a 1-1 correspondence between $f \in F$ and *linear* definitions of Lim, i.e., definitions of Lim for which

$$\underset{s\to\infty}{\text{Lim}}\,[ax(s) + by(s)] = a\,\underset{s\to\infty}{\text{Lim}}\, x(s) + b\,\underset{s\to\infty}{\text{Lim}}\, y(s) \qquad (a, b \in R;\, x, y \in E). \tag{9.01}$$

Some additional properties which a given definition of Lim $x(s)$ may satisfy (or may fail to satisfy) are

$$\underset{-\infty<s<\infty}{\text{g.l.b.}}\; x(s) \leqq \underset{s\to\infty}{\text{Lim}}\, x(s) \leqq \underset{-\infty<s<\infty}{\text{l.u.b.}}\; x(s) \qquad (x \in E), \tag{9.02}$$

$$\varliminf_{s\to\infty} x(s) \leqq \underset{s\to\infty}{\text{Lim}}\, x(s) \leqq \varlimsup_{s\to\infty} x(s) \qquad (x \in E), \tag{9.03}$$

$$\varliminf_{h\to\infty} \frac{1}{h} {}_*\!\int_0^h x(s)\, ds \leqq \underset{s\to\infty}{\text{Lim}}\, x(s) \leqq \varlimsup_{h\to\infty} \frac{1}{h} {}^*\!\!\int_0^h x(s)\, ds \qquad (x \in E), \tag{9.04}$$

$$\underset{s\to\infty}{\text{Lim}}\, x(s + \lambda) = \underset{s\to\infty}{\text{Lim}}\, x(s) \qquad (\lambda \in R;\, x \in E), \tag{9.05}$$

$$\underset{s\to\infty}{\text{Lim}}\, x(\mu s + \lambda) = \underset{s\to\infty}{\text{Lim}}\, x(s) \qquad (\mu > 0;\, \lambda \in R;\, x \in E), \tag{9.06}$$

$$\underset{s\to\infty}{\text{Lim}}\, x(\mu s^\gamma) = \underset{s\to\infty}{\text{Lim}}\, x(s) \qquad (\mu, \gamma > 0,\, x \in E), \tag{9.07}$$

$$\underset{s\to\infty}{\text{Lim}}\, x\left(s + \frac{1}{12}\sin 2\pi s\right) = \underset{s\to\infty}{\text{Lim}}\, x(s) \qquad (x \in E), \tag{9.08}$$

$$\underset{s\to\infty}{\text{Lim}}\, x(\mu s^\gamma + \lambda) = \underset{s\to\infty}{\text{Lim}}\, x(s) \qquad (\mu, \gamma > 0;\, \lambda \in R;\, x \in E), \tag{9.09}$$

where in (9.04), ${}_*\!\int$ and ${}^*\!\!\int$ denote respectively lower and upper Lebesgue integrals.

We leave it to the reader to see what determinations of p_0 and G in Theorem 8.4 give the following applications.

9.1. *The greatest $p \epsilon P$ enforcing* (9.02) *is*

$$p_1(x) = \operatorname*{l.u.b.}_{-\infty<s<\infty} x(s). \tag{9.11}$$

9.2. *The greatest $p \epsilon P$ enforcing* (9.03) *is*

$$p_2(x) = \overline{\lim_{s\to\infty}}\, x(s). \tag{9.21}$$

9.3. *The greatest $p \epsilon P$ enforcing* (9.03) *and* (9.05) *is*

$$p_3(x) = \operatorname*{g.l.b.}_{n>0;\, \lambda_k \epsilon R}\ \overline{\lim_{s\to\infty}}\, \frac{1}{n}\sum_{k=1}^{n} x(s+\lambda_k). \tag{9.31}$$

9.4. *The greatest $p \epsilon P$ enforcing* (9.03) *and* (9.06) *is*

$$p_4(x) = \operatorname*{g.l.b.}_{n>0;\, \mu_k>0;\, \lambda_k \epsilon R}\ \overline{\lim_{s\to\infty}}\, \frac{1}{n}\sum_{k=1}^{n} x(\mu_k s+\lambda_k). \tag{9.41}$$

9.5. *The greatest $p \epsilon P$ enforcing* (9.04) *and* (9.06) *is*

$$p_5(x) = \operatorname*{g.l.b.}_{n>0;\, \mu_k>0;\, \lambda_k \epsilon R}\ \overline{\lim_{h\to\infty}}\, \frac{1}{h}\,{}^*\!\!\int_0^h \frac{1}{n}\sum_{k=1}^{n} x(\mu_k s+\lambda_k)\, ds. \tag{9.51}$$

9.6. *The greatest $p \epsilon P$ enforcing* (9.03) *and* (9.07) *is*

$$p_6(x) = \operatorname*{g.l.b.}_{n>0;\, \mu_k>0;\, \gamma_k>0}\ \overline{\lim_{s\to\infty}}\, \frac{1}{n}\sum_{k=1}^{n} x(\mu_k s^{\gamma_k}). \tag{9.61}$$

It is easy to give many applications of this sort, and the reader may formulate further applications to show even more vividly than the list we have given that in a sense the greater p-functions enforce fewer properties while the smaller p-functions enforce more properties.

It is of interest to observe that the p-function $p_3(x)$ of 9.3, which is the same as $p_B(x)$ of (1.23) used by Banach to establish existence of a linear definition of Lim satisfying (9.03) and (9.05), is the greatest $p \epsilon P$ which enforces these properties. There are in fact many different linear definitions of Lim having properties (9.03) and (9.05), and (by 9.3 and Definition 4.13) p_3 enforces only those properties common to all of them. An example of $x_0 \epsilon E$ for which it is easy to show $p_3x_0 > p_4x_0$ is $x_0(s) = \sin \log^+ s$, where $\log^+ s = 0$ or $\log s$ according as $s \leqq 1$ or $s > 1$. The fact that, for each $\epsilon > 0$, $x_0(s) > 1 - \epsilon$ over arbitrarily large intervals as $s \to \infty$ implies that $p_3x_0 = 1$. The fact that $x_0(e^{\pi}s) + x_0(s) = 0$ for $s > 1$ implies that $p_4x_0 = 0$. This shows that p_3 does not enforce (9.06) or, in other words, that not all linear definitions of Lim having properties (9.03) and (9.05) have property (9.06). On the other hand, property (9.06) obviously implies property (9.05).

We illustrate a significance of the notion of "greatest $p \epsilon P$ which enforces . . ."

by an example. Since $p_3(\sin \log^+ s) = 1$, there exists a linear functional p_7 and a corresponding linear definition of Lim with

$$(9.71)\qquad p_7(x) \equiv \operatorname*{Lim}_{s\to\infty} x(s) \leqq p_3(x) \qquad (x \in E),$$

and

$$(9.72)\qquad p_7(\sin \log^+ s) = \operatorname*{Lim}_{s\to\infty} \sin \log^+ s = 1.$$

Since $p_7 \in F$, we have also $p_7 \in P$, and it follows from (9.71) and (9.3) that p_7 enforces (9.03) and (9.05). Thus p_3 is the unique greatest $p \in P$ which enforces (9.03) and (9.05), while p_7 is merely one of the many $p \in P$ which enforce (9.03) and (9.05). We saw in the preceding paragraph that p_3 is noncommital in regard to (9.06), i.e., enforces neither validity nor failure of (9.06). However, on account of (9.72) and $p_7 \in F$, p_7 actually enforces failure of (9.06).

It was shown in §5 that no linear definition of Lim exists satisfying (9.03), (9.05), and (9.08); hence no $p \in P$ exists which enforces these properties.

It appears to be unknown whether a linear definition of Lim exists satisfying (9.03) and (9.09). If such a definition exists then (Theorems 1.3 and 4.4) $\bar{p} \in P$ exists with

$$(9.8)\qquad \bar{p}(x) \leqq \bar{R}(x) \equiv \operatorname*{g.l.b.}_{n>0;\,\mu_k,\gamma_k>0;\,\lambda_k \in R}\ \overline{\lim_{s\to\infty}}\ \frac{1}{n}\sum_{k=1}^{n} x(\mu_k s^{\gamma_k} + \lambda_k),$$

so that $\bar{R}(x)$ is an r-function. On the other hand, if $\bar{R}(x)$ is an r-function, then $\bar{p} \in P$ exists satisfying (9.8). Since obviously $\bar{R}(x) \leqq p_2(x)$, $p_4(x)$ and $p_6(x)$ it follows that $\bar{p}(x) \leqq p_2(x)$, $p_4(x)$, and $p_6(x)$ so that $\bar{p}$ enforces (9.03), (9.06), and (9.07). But (9.06) and (9.07) imply (9.09). Thus we have shown that a linear definition of Lim exists satisfying (9.03) and (9.09) if and only if $\bar{R}$ is an r-function.

10. **Application to integration.** Let E be the linear space of real functions $x \equiv x(s)$, defined over $-\infty < s < \infty$, for which the upper Lebesgue integral $^*\!\int |x(s)|\,ds$ of $|x(s)|$ over $-\infty < s < \infty$ is finite. Let

$$(10.01)\qquad p_0(x) = {}^*\!\!\int x(s)\,ds \qquad (x \in E).$$

Let G denote the group of transformations g such that if $g \in G$, then real μ and λ with $\mu \neq 0$ exist such that

$$(10.02)\qquad gx(s) = |\mu|\, x(\mu s + \lambda) \qquad (-\infty < s < \infty).$$

It can be verified that $g \in G$ is linear, G is solvable, and $pgx = px$ for $g \in G$, $x \in E$.

Writing $\int_{-\infty}^{\infty} x(s)ds$ in place of $f(x)$ for $x \in E$, we obtain a 1–1 correspondence between $f \in F$ and *linear* definitions of integral, i.e., integrals for which

$$(10.03)\qquad \int_{-\infty}^{\infty} [ax(s) + by(s)]\,ds = a\int_{-\infty}^{\infty} x(s)\,ds + b\int_{-\infty}^{\infty} y(s)\,ds$$

when $a, b \in R$ and $x, y \in E$, and obtain the following application of Theorem 8.4.

10.1. *The greatest* $p \in P$ *enforcing the properties*

$$(10.11) \qquad {}_*\!\int x(s)\,ds \leqq \int_{-\infty}^{\infty} x(s)\,ds \leqq {}^*\!\!\int x(s)\,ds \qquad (x \in E),$$

$$(10.12) \qquad \int_{-\infty}^{\infty} x(\mu s + \lambda)\,ds = \frac{1}{|\mu|}\int_{-\infty}^{\infty} x(s)\,ds \qquad (\mu \neq 0; \lambda \in R; x \in E)$$

is

$$p(x) = \underset{n>0;\,\mu_k>0;\,\lambda_k \in R}{\text{g.l.b.}} {}^*\!\!\int \frac{1}{n}\sum_{k=1}^{n} |\mu_k|\, x(\mu_k s + \lambda_k)\,ds.$$

It is easy to make other specializations of p_0 and G to obtain other applications, some with greater p-functions which enforce fewer properties and some with smaller p-functions which enforce more properties.

CORNELL UNIVERSITY.

THE MEASURE OF TRANSITIVE GEODESICS ON CERTAIN THREE-DIMENSIONAL MANIFOLDS

By Annita Tuller

Introduction. The problem of the existence of transitive geodesics on two-dimensional manifolds of constant negative curvature has been completely solved by Koebe [1].[1] These manifolds are obtained by assigning a hyperbolic metric to the interior of the unit circle in the complex plane and by considering as identical the points congruent under a Fuchsian group.

The question of the measure of the transitive geodesics on such manifolds has also been treated but has not been completely solved. Using the theory of continued fractions, Artin [2] proved that almost all geodesics are transitive if the group is the modular group. Myrberg [3] proved that the same result holds if the group is of the first kind (i.e., one which ceases to be properly discontinuous on the unit circle U), has a finite set of generators and has a fundamental region either lying, with its boundary, entirely inside U or having all its vertices on U. These results are included in the work of E. Hopf [4], who shows that metrical transitivity holds if the group is of the first kind with a finite set of generators. Metrical transitivity implies that almost all the geodesics are transitive. The case of an infinite set of generators has not been considered.

Three-dimensional manifolds of constant negative curvature can be obtained by assigning a hyperbolic metric

$$ds^2 = \frac{4(dx^2 + dy^2 + dz^2)}{(1 - x^2 - y^2 - z^2)^2}$$

to the interior of the unit sphere S and considering as identical the points congruent to each other under suitable groups of the rigid motions of this space. The groups must be properly discontinuous within S but may or may not be properly discontinuous on S. An example of such a manifold is the one obtained by using the Picard group. Recently Löbell [5] gave examples of closed manifolds of constant negative curvature.

As to the properties of geodesics on these manifolds Löbell [5] states that, by methods analogous to those in his proofs for two-dimensional manifolds [6], it can be shown that the periodic geodesics are everywhere dense among the totality of geodesics and that there exist transitive geodesics on the manifolds which he has set up.

It is the object of this paper to prove that, for certain three-dimensional

Received September 29, 1937.

[1] The numbers in brackets refer to the bibliography at the end of the paper.

manifolds of constant negative curvature, almost all geodesics are transitive. For transitivity, as in the two-dimensional case, it is necessary to assume that the group defining the manifold is of the first kind; i.e., that the group is not properly discontinuous in any domain on S. It will then be shown that under this hypothesis the periodic geodesics are everywhere dense among the totality of geodesics. The class of manifolds for which it will be proved that almost all geodesics are transitive is defined by means of a stability property. Hopf [7] has shown that for certain dynamical systems, including those considered in this paper, almost all motions are either stable or unstable—stable in the sense that the motion returns arbitrarily close to any one of the positions formerly occupied, unstable in the sense that it stays outside any finite domain after a finite length of time. It will be proved that, if almost all geodesics are stable, almost all of them are transitive. This includes the set of manifolds defined by groups of the first kind with a finite set of generators. Whether it includes manifolds defined by groups of the first kind with an infinite set of generators is unknown.

The method of proof used in obtaining these results is direct, entirely geometrical and applies equally well to the case of two dimensions.

1. **Three-dimensional hyperbolic geometry.** This section gives a brief summary of the known properties of three-dimensional hyperbolic space which form the background for this paper. The proofs are omitted and references for the material used are given.

Consider the Riemannian space of constant negative curvature $k = -1$, given by the metric

$$ds^2 = \frac{dx^2 + dy^2 + dz^2}{z^2} \tag{a}$$

defining a geometry in the upper half-space $z > 0$. The angle between two curves is the ordinary Euclidean angle. The geodesics are circles orthogonal to the plane $z = 0$ [Bianchi, 8, p. 584]. The geodesic surfaces are spheres orthogonal to the plane $z = 0$.

The rigid motions of the geometry defined must transform the plane $z = 0$ into itself and also must transform two intersecting spheres orthogonal to $z = 0$ into two spheres orthogonal to $z = 0$ and intersecting each other at the same angle. Hence, the transformation thus determined on the points of $z = 0$ must be a conformal transformation. If we introduce into the plane $z = 0$ the complex variable $\zeta = x + iy$, the desired transformation is the linear fractional transformation in this plane [Bianchi, 8, p. 586]:

$$\zeta' = \frac{a\zeta + b}{c\zeta + d}, \qquad ad - bc \neq 0.$$

(The class of anti-analytic transformations

$$\zeta' = \frac{a\bar{\zeta} + b}{c\bar{\zeta} + d}, \qquad ad - bc \neq 0, \qquad \bar{\zeta} = x - iy$$

will not be considered in this paper.) From this the analytic form for the rigid motions in the space itself is very easily developed [Bianchi, 8, pp. 587–588]. It assumes the following form:

$$\rho'^2 = \frac{a\bar{a}\rho^2 + a\bar{b}\zeta + \bar{a}b\bar{\zeta} + b\bar{b}}{c\bar{c}\rho^2 + c\bar{d}\zeta + \bar{c}d\bar{\zeta} + d\bar{d}}, \qquad \rho^2 = x^2 + y^2 + z^2,$$

$$\zeta' = \frac{a\bar{c}\rho^2 + a\bar{d}\zeta + b\bar{c}\bar{\zeta} + b\bar{d}}{c\bar{c}\rho^2 + c\bar{d}\zeta + \bar{c}d\bar{\zeta} + d\bar{d}},$$

$$\bar{\zeta}' = \frac{\bar{a}c\rho^2 + \bar{a}d\bar{\zeta} + \bar{b}c\zeta + \bar{b}d}{c\bar{c}\rho^2 + c\bar{d}\zeta + \bar{c}d\bar{\zeta} + d\bar{d}}.$$

Therefore, corresponding to every rigid motion of the space there is a linear fractional transformation of the plane $z = 0$ into itself, and conversely. A study of this geometry and of the invariant curves under the different types of linear fractional transformations is given by Bianchi [8, pp. 588–589].

We will, however, transform this hyperbolic space with the metric (a) into the space with the metric

$$\text{(b)} \qquad ds^2 = \frac{4(dx^2 + dy^2 + dz^2)}{(1 - x^2 - y^2 - z^2)^2}$$

by transforming the space (x, y, z) into itself so that the upper half-space is transformed into the interior of the unit sphere. This is done by performing in succession the translation

$$x' = x, \qquad y' = y, \qquad z' = z + 2,$$

the inversion

$$r' = \frac{4}{r}, \qquad \theta' = \theta, \qquad \psi' = \psi, \qquad \text{where} \begin{cases} x = r \sin\theta \cos\psi, \\ y = r \sin\theta \sin\psi, \\ z = r\cos\theta, \end{cases}$$

and the translation

$$x' = x, \qquad y' = y, \qquad z' = z - 1.$$

Thus it will suffice for the purpose of this paper to give a brief description of the geometry defined by this new metric and the invariant curves under the linear fractional transformations in the new hyperbolic space.

The metric (b) defines a hyperbolic geometry in the interior of the unit sphere S. The geodesics, or hyperbolic lines, are now arcs of circles orthogonal to S; the hyperbolic planes are parts of spheres orthogonal to S. The equidistant surfaces, i.e., the locus of points at a given hyperbolic distance from a given hyperbolic line, will be called hyperbolic cylinders. The hyperbolic spheres, i.e., the locus of points at a given distance from a given point, are also Euclidean spheres not necessarily with the same center. The rigid motions

are given by the four types of linear fractional transformations. These transform circles and spheres into circles and spheres respectively and preserve angles.

The hyperbolic transformation has two fixed points on the sphere. The geodesic through these points is the axis of the transformation. The family of fixed circles on S determines a family of fixed spheres through the two fixed points, orthogonal to S, and cutting S in the family of fixed circles. The invariant curves of the transformation are arcs of circles through the two fixed points. The transforms of a point P under the iterates of a particular hyperbolic transformation have as limit point one of the fixed points of that transformation. The hyperbolic transformation whose axis is AB and which sends points toward B will be denoted by T_{AB}. The hyperbolic transformation which sends points toward A will be denoted by T_{BA}.

The elliptic transformation also has two fixed points on S and an axis which is the geodesic through these two points. The family of fixed circles on S determines a family of non-intersecting fixed spheres orthogonal to S. The axis is fixed point for point. The invariant curves are circles of intersection of the fixed spheres with the hyperbolic cylinders around the axis. This transformation may be thought of as a rotation about the axis. The transforms of a point P under the iterates of a particular elliptic transformation form a discrete set of points on the invariant circle through P or are everywhere dense on that circle according as the angle of rotation is or is not commensurable with 2π.

The loxodromic transformation is the product of a hyperbolic transformation and an elliptic transformation. The invariant curves wind around the hyperbolic cylinders and cut the circles through the fixed points at a constant angle not equal to zero or $\pi/2$. The transforms of a point P under the iterates of a given loxodromic transformation have as limit point one of the fixed points of that transformation. A notation similar to that used for hyperbolic transformations will be used to show which fixed point is being approached.

The parabolic transformation has one fixed point on S and a family of fixed spheres orthogonal to S, tangent to each other at the fixed point and cutting S in the family of fixed circles. The horospheres, i.e., the Euclidean spheres internally tangent to S at the fixed point, can also be shown to be fixed. The invariant curves are the circles of intersection of the fixed spheres with the horospheres. The transforms of a point under the iterates of a given parabolic transformation have as limit point the fixed point of the transformation.

2. **Groups of linear fractional transformations.** Let G be a group of linear fractional transformations,

$$\zeta' = \frac{a\zeta + b}{c\zeta + d}, \qquad ad - bc = 1.$$

Since G defines an isomorphic group of motions of S and its interior into itself, the notation G can be used equally well for this second group. Two points

P and P' in or on S are said to be *congruent* or *equivalent* if there is a transformation T of G such that $T(P) = P'$. Either point will be called a *copy* of the other.

DEFINITION. The group G will be called *discrete* if the unit element of G is not a point of accumulation of other elements of G [van der Waerden, 9, p. 28]. By this we shall mean that there exists a real positive number α such that for all transformations of G, other than the identity,

$$| a - d |^2 + | b |^2 + | c |^2 > \alpha.$$

Geometrically this means that given a sufficiently small positive ϵ, there exists no transformation T in G, other than the identity, such that the Euclidean distance between P and $T(P)$ is less than ϵ for all P in or on S.

A theorem on discrete groups, which will be used later, is the following:

THEOREM 2.1. *If there exists a positive number a such that, given any positive ϵ, there exists in G two elliptic transformations with fixed points AB and $A'B'$, respectively, such that the distances AB and $A'B'$ are each greater than a, but the distances AA' and BB' are each less than ϵ, then G is not discrete* [F-K, 10, p. 98].

DEFINITION. The group G will be called *properly discontinuous* in a domain D if each point P of D possesses a neighborhood $N(P)$ which has points in common with only a finite number of copies of $N(P)$, and if any two non-congruent points P and Q possess neighborhoods $N(P)$ and $N(Q)$ which are such that no point of $N(P)$ is congruent to a point of $N(Q)$.

DEFINITION. A *fundamental region* of a group G in a domain D is an open set having no points in common with any of its copies and which contains in its interior or on its boundary a point congruent to each point P in D [van der Waerden, 9, p. 35].

If a group is properly discontinuous in a domain D, it possesses a fundamental region in that domain [van der Waerden, 9, p. 35]. It has also been shown that a necessary and sufficient condition that a group of linear fractional transformations be properly discontinuous in S is that it be discrete [F-K, 10, p. 98].

DEFINITION. A group of linear fractional transformations will be said to be of the *first kind* if it is properly discontinuous inside S but not properly discontinuous in any domain on S.

THEOREM 2.2. *The set of transformations in a group of the first kind has fixed points which are everywhere dense on S.*

Since the group is properly discontinuous in S, it is discrete and the Euclidean distance between two congruent points, P and $T(P)$, for generic P (i.e., for P not in the neighborhood of fixed points of T), is greater than a suitably chosen positive number ϵ. Since the group is not properly discontinuous in any domain on S, any neighborhood on S contains at least one pair of congruent points; i.e., we can find a point P and a transformation T such that the Euclidean distance between P and $T(P)$ is less than ϵ. However, it is geometrically evident that this can happen only in the neighborhood of a fixed point. (See F-K, 10, pp. 94–98, for detailed description of the linear fractional transformations of a discrete group.)

The groups with which this paper deals are groups of the first kind. Such a group has a "normal" fundamental region which consists of a simply connected polyhedron bounded by hyperbolic planes, congruent in pairs under transformations of the group. The transformations relating pairs of congruent faces of the normal fundamental region generate the group. The normal fundamental region and its copies cover the interior of S. These and further properties of properly discontinuous groups and their fundamental regions are dealt with at length in [F-K 10, pp. 94–150].

If points in S congruent under the transformations of a group G of the first kind are considered identical, a three-dimensional manifold M is defined. It is the behavior of the geodesics on this manifold that will be investigated.

3. **Periodic geodesics.**

Definition. A directed geodesic in S is *periodic* if it passes through two congruent points with congruent directions.

An immediate consequence of this is that the axis of a hyperbolic or loxodromic transformation is periodic.

It will be shown that, if G is of the first kind, the periodic geodesics are everywhere dense among the totality of geodesics. The following lemma must first be proved.

Lemma 3.1. *Let I_1 and I_2 be two arbitrary circular regions on S having no points in common and bounded by circles O_1 and O_2 respectively. If T is a linear fractional transformation such that $T(O_1) = O_2$ and such that the part of the surface of S exterior to I_1 is transformed by T into the interior of I_2, then T is a hyperbolic or loxodromic transformation with one fixed point in I_1 and one in I_2.*

Since I_1 and I_2 have no point in common, the part of the surface of S exterior to I_1 contains I_2, and hence T transforms a circular region on S into part of itself. We shall show that any such transformation must have a fixed point in I_2.

If T is an elliptic transformation, then, no matter where the fixed points are, there are fixed circles either cutting one of the bounding circles and not the other or lying entirely in the region between them. Neither case is possible under the hypothesis of the lemma. Therefore T is not elliptic.

If T is parabolic, hyperbolic or loxodromic, the transforms of a point P_1 on O_1 under the iterates of T have as limit point a fixed point of T. However, $T(P_1) = P_2$ on O_2. Therefore, further transforms of P_1 under powers of T are inside I_2, and thus T has a fixed point inside I_2. A similar argument applied to T^{-1} gives us a fixed point inside I_1. Hence T is either a hyperbolic or a loxodromic transformation of the type desired.

Theorem 3.1. *If G is of the first kind and if I_1 and I_2 are arbitrary circular neighborhoods on S with no point in common, there exists a hyperbolic or loxodromic transformation of G with one fixed point in I_1 and the other in I_2.*

It is convenient to divide the proof into three parts.

Case 1. *The group G contains a parabolic transformation T_Q but no elliptic transformation.*

It will first be shown that a hyperbolic or loxodromic transformation of G cannot have a fixed point in common with a parabolic one. If Q is a fixed point of a hyperbolic or loxodromic transformation whose axis is AQ, and if Q is also the fixed point of a parabolic transformation T_Q, then $T_Q(AQ) = A'Q$ is the axis of a hyperbolic or loxodromic transformation. We now have two hyperbolic or loxodromic transformations with a common fixed point. This, however, will be shown to lead to a contradiction. By applying powers of $T_{A'Q}$ to a point P on the axis $A'Q$ we get congruent points arbitrarily close, in the hyperbolic sense, to the hyperbolic line AQ. We now construct a sphere S_1, orthogonal to S and to AQ. Let $S_2 = T_{QA}(S_1)$. By applying sufficiently high powers of T_{QA}, and by choosing a proper subset, we get an infinite sequence of distinct points congruent to P and having as limit point a point on the part of AQ between S_1 and S_2. However, this is impossible for a properly discontinuous group.

It will now be shown that the fixed points of parabolic transformations are everywhere dense on S. Let I be an arbitrary circular region on S. There is a fixed point A in it (Theorem 2.2). If the corresponding transformation is not parabolic, then it is hyperbolic or loxodromic and its other fixed point B is not Q. By applying sufficiently high powers of T_{BA}, we get Q', congruent to Q, inside I. Since Q' is the fixed point of a parabolic transformation and I is any region on S, we have the parabolic fixed points everywhere dense on S.

We come now to the theorem itself. In I_1 there is a fixed point P_1 of a parabolic transformation. Let I_3 be a circular region containing all of I_1 but no part of I_2. Let O_3 be the boundary of I_3. By applying sufficiently high powers of T_{P_1} we get O_3 transformed into O_1' entirely in I_1. The part of S exterior to I_3 will be transformed into a region I_1', bounded by O_1', inside I_1. This can be seen by considering the behavior under the transformation of any one point exterior to I_3. In a similar manner we get O_2' congruent to O_3 in I_2. This time, however, the interior of I_3 is transformed into a region I_2', bounded by O_2', inside I_2. We note now that the regions I_1' and I_2' satisfy the conditions of the lemma, and the theorem is proved for this case.

CASE 2. *The group G contains only hyperbolic and loxodromic transformations.*

Since the fixed points are everywhere dense on S (Theorem 2.2), I_1 contains a fixed point C_1 of a hyperbolic or loxodromic transformation whose other fixed point is D_1; and I_2 contains a fixed point C_2 of a hyperbolic or loxodromic transformation whose other fixed point is D_2. Let I_3 be a circular neighborhood containing I_1 but not I_2, and bounded by a circle O_3 which passes through neither D_1 nor D_2. By applying sufficiently high powers of $T_{D_1C_1}$ we get a copy O_1' of O_3 in I_1 bounding a region I_1' in I_1. Similarly, we get a copy O_2' of O_3 in I_2 bounding a region I_2' in I_2. However, I_1' and I_2' may not satisfy the conditions of the lemma, since the interior of I_1' may correspond to the interior of I_2'. We will show, however, that there exists a transformation of the group which takes the exterior of I_2' into a region I_2'' inside I_2'. Then I_1' and I_2'' will satisfy the conditions of the lemma.

Let A and B be the fixed points of a hyperbolic or loxodromic transformation T_{AB}. By taking a region I_4 inside I_2' containing neither A nor B and treating it as we did I_1 above we get copies A' and B' of A and B, respectively, inside I_4. This can be done since, as has been shown under Case 1, two hyperbolic or loxodromic transformations cannot have a common fixed point. By applying sufficiently high powers of $T_{A'B'}$ we get a copy O_2'' of O_2' inside I_2', bounding a region I_2'' in I_2' such that the exterior of I_2' corresponds to the interior of I_2''.

CASE 3. *The group G contains an elliptic transformation.*

In a manner similar to that of the first part of Case 1, it can be shown that an elliptic transformation cannot have a fixed point in common with a hyperbolic or loxodromic transformation unless the whole axis is common.

We go on to show that any neighborhood on S contains a pair of fixed points of an elliptic transformation. Let PQ be the axis of our known elliptic transformation. First, let us suppose that the only fixed points in I_1 are elliptic. Then there must be a pair in I_1 belonging to the same transformation. Otherwise, the correspondents of all the fixed points inside any subneighborhood of I_1 would have a cluster point outside or on the boundary of I_1, and this contradicts Theorem 2.1. Moreover, by applying the same argument to smaller neighborhoods inside I_1, it is possible to obtain a pair of elliptic fixed points arbitrarily close together.

If a fixed point Q' inside I_1 is parabolic, then by applying sufficiently high powers of $T_{Q'}$ we get copies of P and Q inside I_1, again arbitrarily close. If a fixed point inside I_1 is hyperbolic, the other fixed point of the transformation is not P or Q and again we can get copies of P and Q inside I_1. The same is true if the fixed point inside I_1 is loxodromic.

We come now to the theorem itself. Let O_3 be a great circle separating I_1 and I_2. Let E be the midpoint of the elliptic axis P_1Q_1 whose fixed points P_1 and Q_1 are in I_1. Let the plane through E and O perpendicular to the axis P_1Q_1 cut O_3 in points A and B, and the boundary of I_1 in points C and D. Then AB is a diameter of S. Let θ_1, θ_2, θ_3, θ_4 be the angles in rotation about E between the geodesic rays EC and EA, EA and EB, EB and ED, ED and EC respectively.

Since we can get P_1 and Q_1 arbitrarily close, we can get E arbitrarily near the surface, and thus angles θ_1, θ_2, and θ_3 can be made arbitrarily small, while angle θ_4 can be made arbitrarily close to 2π. Therefore, no matter what the rotation angle of our elliptic transformation is, we can get E close enough to the surface of the sphere so that repeated application of the elliptic transformation $T_{P_1Q_1}$ will take the circle O_3 into a circle O_1' in I_1. In a similar manner we get a copy O_2' of O_3 in I_2.

If we let I_3 be the region of S bounded by O_3 and containing I_1, the above process transforms the exterior of I_3 into the interior of the region I_1' in I_1 bounded by O_1'. Since O_3 separates I_1 and I_2, the interior of I_3 will be transformed into the interior of the region I_2' in I_2 bounded by O_2'. I_1' and I_2' satisfy the conditions of the lemma. This completes the proof.

4. Transitive horospheres. Let E be the set of elements

$$p:\ (x, y, z; \psi, \theta),$$

where

$$x^2 + y^2 + z^2 < 1, \qquad 0 \leqq \psi \leqq \pi, \qquad 0 \leqq \theta < 2\pi.$$

Any such element determines a point $P:(x, y, z)$ inside S and a direction $d:(\psi, \theta)$ at this point, where ψ is the angle between d and a line z' through P parallel to the z-axis, and θ is the angle between the plane determined by $z'd$ and the plane through P parallel to the (x, z)-plane. The elements will be briefly designated by $p:(x; d)$. Any such element determines a directed geodesic since there is one and only one geodesic through a point P inside S in a given direction d. We define the distance between two elements $p_1:(x_1; d_1)$ and $p_2:(x_2; d_2)$ as

$$\| p_2 - p_1 \| = H(P_1P_2) + \alpha,$$

where $H(P_1P_2)$ is the hyperbolic distance between the points $P_1:(x_1)$ and $P_2:(x_2)$, and α is the least positive angle between d_1 and d_2.

Definition. Two elements $p_1:(x_1; d_1)$ and $p_2:(x_2; d_2)$ of E are *congruent* if there is a transformation of G taking $P_1:(x_1)$ into $P_2:(x_2)$, and the direction d_1 at P_1 into the direction d_2 at P_2; i.e., taking P_1 into P_2 and the directed geodesic determined by p_1 into the directed geodesic determined by p_2.

Definition. An element of E is *periodic* if it determines a periodic geodesic.

It follows that all the elements on a periodic geodesic are periodic elements.

Theorem 4.1. *If G is of the first kind, the periodic elements are everywhere dense in E.*

Let $N(p)$ be any neighborhood of any element $p:(x; d)$ of E. The element p determines a directed geodesic AB. Let I_A and I_B be sufficiently small regions on S about A and B respectively, so that the set of geodesics having initial endpoint in I_A and terminal endpoint in I_B goes through $N(p)$. However, by Theorem 3.1, there is at least one periodic geodesic in that set and hence a periodic element in $N(p)$.

Definition. A *horosphere* is a Euclidean sphere internally tangent to S.

The point at infinity of a horosphere is its point of contact with S. A horosphere is completely determined by its Euclidean radius and its point at infinity. The horosphere with Euclidean radius r and point at infinity Q will be denoted by $h(Q, r)$. The elements of E on the horosphere will be taken as its points inside S with a direction at each point normal to the horosphere and directed outward. It is clear that an element of E determines a unique horosphere, and that a point A in S and a point B on S determine a unique horosphere through A with point at infinity at B.

We may also note that two horospheres $h(Q, r_1)$ and $h(Q, r_2)$ with Q as point at infinity cut off equal hyperbolic lengths on the set of geodesics through Q. For, let QA_1A_2R and QB_1B_2S be any two geodesics through Q. The points

A_1 and B_1 are on $h(Q, r_1)$, and A_2 and B_2 on $h(Q, r_2)$. We can find a parabolic transformation T_Q such that $T_Q(A_1) = B_1$. Then $T_Q(QR) = QS$. Since the horospheres are fixed under T_Q, $T_Q(A_2) = B_2$. Hence the hyperbolic distances A_1A_2 and B_1B_2 are equal.

DEFINITION. A horosphere $h(Q, r)$ is *transitive* if the totality of elements on it and on all its copies form a set everywhere dense in E.

The remainder of this section is devoted to several theorems on the transitivity of horospheres. Where the proofs are completely analogous to those for horocycles in the two-dimensional hyperbolic space, the exact reference is given at the end of the statement of the theorem. The proofs are omitted.

DEFINITION. A set of horospheres is *transitive* if the totality of elements on all copies of all members of the set is everywhere dense in E.

LEMMA 4.2. *Let P be a point inside S and I a circular domain on S. Let I' be a circular domain of I, and $h(P, I')$ the set of horospheres through P with points at infinity in I'. If $h(P, I')$ is transitive for every circular domain I' of I, then there exists an infinite set of transitive horospheres through P with points at infinity in I* [Hedlund, 11, Lemma 1.1, p. 532].

THEOREM 4.2. *If G is of the first kind and if P is an arbitrary point in S and I an arbitrary circular region on S, there exists a point Q in I such that $h(P, Q)$ is transitive.*

The theorem will follow from the lemma if we show that the set $h(P, I)$ is transitive. Let AB be the axis of a hyperbolic or loxodromic transformation of G, and hence a periodic geodesic. It has a copy $A'B'$ with endpoints in I. Let P_0 be a point on AB. It has a copy P_0' on $A'B'$. By repeated application of $T_{A'B'}$ we get copies $P_1', P_2', \cdots$ of P_0 approaching B'. Consider a set of horospheres, all passing through P and one through each P_n'. As n becomes infinite, the set of horospheres through P and P_n' will have points at infinity Q_n approaching B'. The angle at which the horospheres cut $A'B'$ will approach $\pi/2$. Therefore, if $p_n:(x_n; d_n)$ is the element on $A'B'$ at the point of intersection with the horosphere through P and P_n', and $q_n:(x_n; d_n')$ is the element on the horosphere at the point of intersection with $A'B'$, then, as n becomes infinite, the angle between d_n and d_n' approaches zero, and hence $\| q_n - p_n \|$ approaches zero. Since p_0 is any periodic element, and since the periodic elements are everywhere dense in E, the set $h(P, I)$ is transitive.

THEOREM 4.3. *If G is of the first kind, there exists an infinite set of transitive horospheres through any point in S.*

The points at infinity of these form a set everywhere dense on S. This theorem is an immediate consequence of the preceding one.

THEOREM 4.4. *If one horosphere with Q as point at infinity is transitive, all the horospheres with Q as point at infinity are transitive* [Hedlund, 11, Theorem 2.1, p. 535].

DEFINITION. A point Q of S will be called *h-transitive* if all the horospheres with Q as point at infinity are transitive.

THEOREM 4.5. *If G is of the first kind, the endpoints of all axes of hyperbolic*

or loxodromic transformations of G are h-transitive [Hedlund, 11, Theorem 2.2, p. 536].

THEOREM 4.6. *If G is of the first kind and there are copies of the horosphere $h(Q, r)$ with radii arbitrarily close to* 1, *Q is h-transitive* [Hedlund, 11, Theorem 2.3, p. 537].

THEOREM 4.7. *If G is of the first kind, Q a point on S, and if there exists on OQ a sequence of points $O_1, O_2, \cdots$ such that the limit as n becomes infinite of the hyperbolic distance OO_n is $+\infty$ and such that O_n has a copy O'_n with hyperbolic distance OO'_n bounded, then Q is h-transitive* [Hedlund, 11, Theorem 2.4, p. 537].

THEOREM 4.8. *If G is of the first kind with a fundamental region R_0 which together with its boundary lies entirely inside S, then all points on S are h-transitive.*

This is an immediate consequence of the preceding theorem.

5. The measure of transitive geodesics.

DEFINITION. A directed *geodesic* is *transitive* if the totality of elements on it and on all its copies form a set everywhere dense in E.

DEFINITION. An *element* of E will be called *transitive* if it determines a transitive geodesic.

With the aid of the preceding theory on transitive horospheres we derive some theorems concerning the measure of the transitive geodesics on the manifold M, defined by considering as identical the points in S congruent under the transformations of a group G of the first kind. Measure in E is Lebesgue measure but the space will be considered as a Euclidean space.

THEOREM 5.1. *The transitive elements of E form a measurable set.*

Let $\epsilon_1, \epsilon_2, \cdots$ be a sequence of positive numbers such that $\lim_{n\to\infty} \epsilon_n = 0$. Let E be covered by a finite set of open cells such that the Euclidean diameter of each cell is less than ϵ_1 and such that every cell contains some of E. Let O_1 be the set of elements of E, each of which determines a geodesic such that every cell contains an element of it or of one of its copies. The set O_1 is open; for, if p_1 is an element of O_1 and p_2 is an element in a sufficiently small neighborhood of p_1, the geodesic g_2 determined by p_2 will stay close enough to that determined by p_1 so that every one of our open cells will contain an element of g_2 or of one of its copies. Let E be now covered by a finite set of open cells each of diameter less than ϵ_2 and such that every cell contains some of E; and let O_2 be the set of elements, each of which determines a geodesic such that every one of these new cells contains an element of that geodesic or one of its copies. Continuing in this manner we get sets $O_1, O_2, O_3, \cdots$. An element is transitive if and only if it belongs to all of these sets. Hence the set of transitive elements is $\prod_{n=1}^{\infty} O_n$ which is measurable since each O_i is open and therefore measurable.

THEOREM 5.2. *If G is of the first kind with a fundamental region R_0 which, together with its boundary, lies entirely within S, any measurable set W in E, where W is of positive measure, determines a transitive set of geodesics.*

We shall consider first a special set Φ, the set determining a solid cone of geodesic rays with vertex at the point P in S of element $p:(x;\bar{d})$ and with generators determined by the cone of directions through P making an angle less than or equal to ϕ with $\bar{d}$. We shall show that the totality of elements on all the geodesic rays determined by Φ is a transitive set.

On each ray of the cone consider the element $(x_s\,;d_s)$ of E whose point (x_s) in S is at a hyperbolic distance s from P. The totality of such points (x_s) is on a spherical zone a_s of hyperbolic area A_s. The elements $(x_s\,;d_s)$ have directions normal to a_s. As s becomes infinite the hyperbolic radius of the zone becomes infinite. Let s become infinite through a sequence of values $s_1 < s_2 < \cdots$ and consider the elements $(x_{sn}\,;\bar{d}_{sn})$ thus determined on the central geodesic ray g_c. Each such element has a copy $(x^*_{sn};\bar{d}^*_{sn})$ in R_0. Since we can pick a subsequence of these copies with a limiting element, we shall assume that this is already the subsequence with unique limiting element $(x^*;\bar{d}^*)$ in R_0. This element determines a horosphere $h(Q, r)$ which is transitive (Theorem 4.8). Let T_n be the transformation of G which carries $(x_{sn}\,;\bar{d}_{sn})$ into $(x^*_{sn};\bar{d}^*_{sn})$, and let a^*_{sn} be $T_n(a_{sn})$. As n becomes infinite a^*_{sn} is a sequence of spherical zones whose hyperbolic radii become infinite. Each a^*_{sn} carries an element e^*_{sn} at its center such that the $\lim_{n\to\infty} e^*_{sn}$ is $(x^*;\bar{d}^*)$ and such that a^*_{sn} approximates more and more of the transitive horosphere $h(Q, r)$. Hence the limiting position of a^*_{sn} is the transitive horosphere $h(Q, r)$. Now let e be any element of E, and let $N(e)$ be a neighborhood of this element. Since $h(Q, r)$ is transitive, there exists an element e_h on $h(Q, r)$ which has a copy in $N(e)$. Let $N(e_h)$ be a neighborhood of e_h so small that each element in $N(e_h)$ has a copy in $N(e)$. If we let Φ_n denote the set $(x_{sn};d_{sn})$, where the points (x_{sn}) are on the spherical zone a_{sn}, there exists an N such that for n greater than N the set Φ_n has an element with a copy in $N(e_h)$ and hence in $N(e)$. Thus the totality of elements on all the geodesic rays determined by Φ is a transitive set.

We now return to the arbitrary set W in E, where W is of positive five-dimensional measure. We take sections of it by planes determined by fixing particular (x, y, z). There exists on at least one of these planes π a section of W of positive plane measure [Hobson, 12, p. 185], and there is an element $p:(x;\bar{d})$ of this section at which the two-dimensional metrical density is 1 [Hobson, 12, p. 179]. This element determines a geodesic ray g_r, and letting s become infinite through a sequence of values $s_1 < s_2 < \cdots$, we again consider the elements $(x_{sn}\,;d_{sn})$ on g_r. Let C_n be the hyperbolic sphere with center x and radius s_n. Given a positive constant A, let a_n be a spherical zone of hyperbolic area A, on C_n and with center at x_{sn}. The set of elements φ_n at the point x and determining rays passing through a_n forms a circular region in π with center at $p:(x;\bar{d})$. As n becomes infinite, the radius of the set φ_n approaches zero, and we have

$$\lim_{n\to\infty}\frac{m(\varphi_n\cdot W)}{m(\varphi_n)} = 1, \tag{1}$$

where the measures are two-dimensional measures. (The cells used in the definition of metrical density are squares; but it is easily seen that if the metrical density is 1 when squares are employed, the same must be true when circles are used.)

Let Φ_n be the elements of the geodesic rays, with initial point x and passing through a_n, at the points where they intersect a_n. Let W_n be the subset of Φ_n obtained when only rays determined by the set $\varphi_n \cdot W$ are considered, and let w_n be the set of points on a_n bearing the elements W_n. The set w_n is a measurable set. We wish to show that

$$\lim_{n \to \infty} \frac{m(w_n)}{A} = 1,$$

where the measure here is in terms of hyperbolic area on a_n.

It can be assumed that $p:(x; \bar{d})$ is such that x is at the center O of S, and $\bar{d}$ is such that its ψ is neither 0 nor π. For if the point x is transformed to O by a rigid motion of the hyperbolic space, angles are preserved and hyperbolic areas are preserved. The transformation can evidently be chosen in infinitely many ways such that the condition on direction is fulfilled. Then for each n, A is the hyperbolic area of a zone of a sphere whose equation in spherical coordinates is

$$x = r_n \sin \psi \cos \theta,$$

$$y = r_n \sin \psi \sin \theta,$$

$$z = r_n \cos \psi,$$

where $0 < r_n < 1$. The hyperbolic metric on this sphere becomes

$$ds^2 = \frac{r_n^2}{(1 - r_n)^2} (\sin^2 \psi \, d\theta^2 + d\psi^2),$$

and we have

$$m(w_n) = \int\int_{\varphi_n \cdot W} \frac{r_n^2}{(1 - r_n)^2} \sin \psi \, d\theta \, d\psi,$$

while

$$A = \int\int_{\varphi_n} \frac{r_n^2}{(1 - r_n)^2} \sin \psi \, d\theta \, d\psi,$$

where both integrals are Lebesgue integrals. Thus

$$\frac{m(w_n)}{A} = \frac{\displaystyle\int\int_{\varphi_n \cdot W} \sin \psi \, d\theta \, d\psi}{\displaystyle\int\int_{\varphi_n} \sin \psi \, d\theta \, d\psi}.$$

Now

$$\frac{m(\varphi_n \cdot W)}{m(\varphi_n)} = \frac{\displaystyle\iint_{\varphi_n \cdot W} d\theta\, d\psi}{\displaystyle\iint_{\varphi_n} d\theta\, d\psi}.$$

Letting M_1 and m_1 be the maximum and minimum, respectively, of $\sin\psi$ in $\varphi_n \cdot W$, and M_2 and m_2 the maximum and minimum, respectively, of $\sin\psi$ in φ_n, we have

$$\frac{m_1 \displaystyle\iint_{\varphi_n \cdot W} d\theta\, d\psi}{M_2 \displaystyle\iint_{\varphi_n \cdot W} d\theta\, d\psi} \leqq \frac{\displaystyle\iint_{\varphi_n \cdot W} \sin\psi\, d\theta\, d\psi}{\displaystyle\iint_{\varphi_n} \sin\psi\, d\theta\, d\psi} \leqq \frac{M_1 \displaystyle\iint_{\varphi_n \cdot W} d\theta\, d\psi}{m_2 \displaystyle\iint_{\varphi_n} d\theta\, d\psi}.$$

As n becomes infinite, ψ approaches a constant different from 0 or π; namely, the ψ of the element $p:(x;\, \bar{d})$; and the first and third ratios become equal. Hence,

$$\lim_{n\to\infty} \frac{m(\varphi_n \cdot W)}{m(\varphi_n)} = \lim_{n\to\infty} \frac{m(w_n)}{A}.$$

From (1),

$$\lim_{n\to\infty} \frac{m(w_n)}{A} = 1. \tag{2}$$

Again let e be any element of E and let $N(e)$ be a neighborhood of this element. If we use the notation of the first case considered under this theorem, we see that there exists an element e_h on $h(Q, r)$ which has a copy in $N(e)$. Let $N(e_h)$ be a neighborhood of e_h so small that each element in $N(e_h)$ has a copy in $N(e)$. Let A be chosen so that the zone a of $h(Q, r)$ with center at x^* and hyperbolic area A contains the point bearing e_h. The sequence of zones $T_n(a_n)$ approaches the zone a uniformly, and for n sufficiently large one of the elements $T_n(W_n)$ will lie in $N(e_h)$. For if this were not the case, there would exist a $\delta > 0$ and an N such that for $n > N$ the zone $T_n(a_n)$ would contain a domain of hyperbolic area greater than δ and containing no point of the set $T_n(w_n)$. The same would hold with respect to the zone a_n and the set w_n, and thus

$$\lim_{n\to\infty} \frac{m(w_n)}{m(a_n)} \leqq 1 - \frac{\delta}{A}.$$

This contradicts (2). Thus the totality of elements on all the geodesic rays determined by W is a transitive set.

THEOREM 5.3. *If G is of the first kind with fundamental region R_0 which,*

together with its boundary, lies entirely within S, then almost all elements of E are transitive.

This theorem will be proved by showing that the complement $C\left(\prod_{n=1}^{\infty} O_n\right)$ of the set $\prod_{n=1}^{\infty} O_n$ of Theorem 5.1 is a zero set. Since $C\left(\prod_{n=1}^{\infty} O_n\right)$ is equal to $C(O_1) + C(O_2) + \cdots$, it will suffice to show that $C(O_i)$ has zero measure. Let the set of open sets in the mesh defined by ϵ_1 be denoted by ${}_1O_1, {}_2O_1, \cdots, {}_nO_1$. Let E_i be the set of elements each of which determines a geodesic such that ${}_iO_1$ contains no element of that geodesic or of any of its copies. Then $C(O_1) = E_1 + E_2 + \cdots + E_n$. The set E_i is measurable since its complement is open and therefore measurable. By Theorem 5.2 each E_i is a zero set. Therefore $C(O_1)$ is a zero set. In a similar manner we show that $C(O_i)$ is a zero set for every i. Hence $C\left(\prod_{n=1}^{\infty} O_n\right)$ is of measure zero.

The proof of Theorem 5.2, and hence that of Theorem 5.3, depends on the group's being such that all points on S are h-transitive. However, we may have groups of the first kind where all points on S are not necessarily h-transitive. For example, if Q is a fixed point of two parabolic transformations of the group, not both in the same cyclic subgroup (e.g., the Picard group), the fundamental region on a horosphere with Q as point at infinity, for the subgroup generated by these two transformations, is of finite hyperbolic area. Hence, the horosphere cannot be transitive. We wish, therefore, to extend Theorem 5.3 to include such groups of the first kind.

Let E_M be the space obtained from E by considering congruent elements identical. This is then the phase space associated with the manifold M. To the uniform motion along the geodesics on M there corresponds a steady flow on E_M. The element of five-dimensional volume $dm = dv\, d\theta\, d\psi$, dv being the three-dimensional hyperbolic element of volume, is invariant under the flow [cf. Hopf, 4, pp. 300–304].

Definition. An element p of E_M will be called *unstable* [Hopf, 7, "fliehende"] if the motion along the geodesic ray determined by it ultimately passes out of and stays out of any finite neighborhood in E_M once occupied. The geodesic ray determined by an unstable element will be called an *unstable geodesic ray*.

Definition. An element p of E_M will be called *stable* [Hopf, 7, "wiederkehr"; semi-stable in the sense of Poisson] if the motion along the geodesic ray determined by it returns infinitely often to an arbitrarily small neighborhood of p. The geodesic ray determined by a stable element will be called a *stable geodesic ray*.

Theorem 5.4. *If G is of the first kind and if the unstable elements form a set of measure zero, almost all geodesics are transitive.*

Each element in E_M determines an infinite set of congruent elements of E. If the element of E_M is stable, each of the corresponding elements $(x; d)$ of E

will determine a geodesic ray which has on it an infinite sequence of elements $(x_{sn}; d_{sn})$ with $\lim_{n\to\infty} s_n = +\infty$ and such that each $(x_{sn}; d_{sn})$ has a copy $(x^*_{sn}; d^*_{sn})$ with $\lim_{n\to\infty}(x^*_{sn}; d^*_{sn}) = (x; d)$. Therefore the geodesic ray ends in a point on S which is h-transitive (Theorem 4.7). It has been shown that, under the hypothesis of this theorem, almost all the elements of E_M are stable [Hopf, 7, Theorem 1, p. 712]. Then almost all the elements of E are correspondents of stable elements of E_M and thus have the property described above.

The proof of this theorem now follows closely that of Theorem 5.2 with the following modifications. In choosing the element $p:(x; \bar{d})$ which is to determine the central geodesic ray g_c we choose one which is stable and such that the element with the same $P:(x)$ but direction opposite to $\bar{d}$ is also stable. This excludes only a set of measure zero; for, since almost all elements of E_M are stable, then, of those elements $p:(x; d)$ which are stable, only a zero set can be such that the elements with the same $P:(x)$ but direction opposite to d are not stable. This assures us that the limiting element $(x^*; d^*)$ is inside S (it is p itself) and that the unique horosphere determined by it is transitive. We do the same in choosing the element $p:(x; \bar{d})$ at which the metrical density is 1. This again excludes only a set of measure zero [Hobson, 12, p. 179]. The rest of this proof and that of Theorem 5.3 go through without change and thus Theorem 5.4 is proved.

THEOREM 5.5. *If G is a group of the first kind with a finite number of generators, all points on S, with the exception of those which are fixed points of more than one cyclic parabolic subgroup of G, are h-transitive.*

A group with a finite number of generators can have only a finite number of points Q_i $(i = 1, \cdots, m)$ of the boundary of the fundamental region R_0 on S. Each of the points Q_i is a fixed point of more than one cyclic parabolic subgroup of the given group G [F-K, 10, pp. 125–126] and hence is not h-transitive. If we let Q be the set of points Q_i $(i = 1, \cdots, m)$ and their copies, this theorem will prove that all points on S except those of set Q are h-transitive.

If the radii r_i $(i = 1, \cdots, m)$ of the horospheres $h(Q_i, r_i)$ are chosen sufficiently near 1, it is geometrically evident that any point of R_0 or its boundary and interior to S will be interior to some one of the set $h(Q_i, r_i)$ $(i = 1, \cdots, m)$. If C denotes the set $h(Q_i, r_i)$ and all copies of these, any point inside S is interior to some member of C.

There exist at least two parabolic transformations of G with fixed point Q_i. Hence each point of $h(Q_i, r_i)$ with the exception of Q_i has a copy within hyperbolic distance D_i of the origin O, where D_i depends on $h(Q_i, r_i)$ and not on the chosen point on it. If D is a constant denoting the largest of the D_i, any point of the set C which is not on S has a copy within hyperbolic distance D of O.

Now let P be any point on S, not belonging to the set Q. If P' is any point on the ray OP, it lies interior to one of the horospheres of C. But $P'P$ cannot lie entirely in any one member of C, for this would imply that P belonged to

the set Q. Hence $P'P$ must intersect one of these horospheres and has on it a point P'' with a copy within hyperbolic distance D of O. From Theorem 4.7, P is h-transitive.

THEOREM 5.6. *If G is a group of the first kind with a finite number of generators, the unstable elements of E_M form a set of measure zero.*

From Theorem 5.5 all points on S not belonging to set Q are h-transitive and hence cannot be endpoints of unstable geodesics. Therefore, the unstable elements of E_M determine geodesics in S ending in the points of set Q. These endpoints form a denumerable set on S. We may have an infinite number of geodesic rays going out to each point of Q. However, almost all unstable geodesic rays are also unstable when extended backward through the initial point [Hopf, 7, Theorem 2, p. 712]. Hence both ends of almost all geodesics determined by unstable elements must belong to the set Q. Therefore, the unstable elements form a set of measure zero.

An immediate consequence of Theorem 5.4 and Theorem 5.6 is the following:

THEOREM 5.7. *If G is of the first kind with a finite number of generators, almost all geodesics are transitive.*

BIBLIOGRAPHY

1. P. KOEBE, *Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen*, fünfte Mitteilung, Sitzungsberichte der Preussischen Akademie der Wissenschaften (1930), pp. 304–364.
2. E. ARTIN, *Ein mechanisches System mit quasiergodischen Bahnen*, Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität (1924), pp. 170–175.
3. P. J. MYRBERG, *Ein Approximationssatz für die fuchsschen Gruppen*, Acta Mathematica, vol. 57 (1931), pp. 389–409.
4. E. HOPF, *Fuchsian groups and the ergodic theory*, Transactions of the American Mathematical Society, vol. 39 (1936), p. 299.
5. F. LÖBELL, *Beispiele geschlossener dreidimensionaler Clifford-Kleinscher Räume negativer Krümmung*, Berichten der mathematisch-physischen Klasse der Sächsischen Akademie der Wissenschaften zu Leipzig (1931), pp. 167–174.
6. F. LÖBELL, *Die überall regulären unbegrenzten Flächen fester Krümmung*, Inaugural-Dissertation, Eberhard-Karls-Universität zu Tübingen (1927).
7. E. HOPF, *Zwei Sätze über den Wahrscheinlichen Verlauf der Bewegungen dynamischer Systeme*, Mathematische Annalen (1930), pp. 710–716.
8. L. BIANCHI, *Vorlesungen über Differentialgeometrie*, Leipzig, 1899.
9. B. L. VAN DER WAERDEN, *Gruppen von Linearen Transformationen*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 4 (1935).
10. R. FRICKE AND F. KLEIN, *Vorlesungen über die Theorie der Automorphen Functionen*, vol. 1, Leipzig, 1897.
11. G. A. HEDLUND, *Fuchsian groups and transitive horocycles*, this Journal, vol. 2 (1936), pp. 530–542.
12. E. W. HOBSON, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 1, Cambridge, 1921.

BRYN MAWR COLLEGE.

CHARACTERIZATION OF THE CONFORMAL GROUP AND THE EQUI-LONG GROUP BY HORN ANGLES

By Edward Kasner

We begin by giving certain preliminary definitions. A *horn angle* consists of two curves which pass through a common point in a common direction. In this paper we consider only horn angles of first order contact, that is, the curves of the horn angle have different curvatures at the common point. By a *contact transformation* we mean a lineal element transformation by which every curve (union) corresponds to a curve. Thus it follows that every contact transformation converts every horn angle into a horn angle. In our previous work[1] we have defined a *natural conformal measure* M_{12} and a distinct *natural equi-long measure* μ_{12} of a horn angle. In this paper we shall determine all contact transformations that preserve M_{12} or μ_{12}. Our results are that the group for which M_{12} is invariant is the direct conformal group; and, dually, the group for which μ_{12} is invariant is the direct equi-long group.

Separate proofs are required for these two results. For, although the two theories are analogous (or roughly dual), they are not connected by any known automatic principle of duality or transference principle. In fact we will notice in the course of our work that some of our preliminary theorems lead to results which are not strictly dual.

We shall consider the geometric properties of certain three-parameter families of curves designated as L-families and λ-families. They may be regarded as generalizations of dynamical trajectories. One of our principal results is that the group of transformations which convert every L-family of curves into an L-family of curves is the conformal group. On the other hand, the group of transformations which convert every λ-family of curves into a λ-family of curves is a group of line transformations which convert parallel lines into parallel lines. This group is much larger than the equi-long group, so that our two results are not dual. The dynamical type group is projective.

Finally we shall prove that there is no contact transformation which changes every conformal measure M_{12} into an equal equi-long measure μ_{12}. This is an

Received October 6, 1937.

[1] Proceedings of the International Congress of Mathematicians, Cambridge, No. 2 (1912), p. 81.
Proceedings of the National Academy of Sciences (with George Comenetz), vol. 22 (1936), p. 303.
Science, vol. 85 (1937), p. 480.
Proceedings of the National Academy of Sciences, vol. 23 (1937), p. 337.
Annals of Mathematics, October, 1937.
Transactions of the American Mathematical Society (1938).

immediate consequence of the theorem that there is no contact transformation which transforms every L-family of curves into a λ-family of curves.

In the first part of our paper we give a discussion of two auxiliary curves: (1) the extended lemniscate, a quartic curve, which is used in characterizing the L'-families of curves; and (2) the extended cissoid, a cubic curve, which is used in characterizing the λ-families of curves. Then we introduce the conformal measure M_{12} and the equi-long measure μ_{12} of a horn angle; and define wide-open trihorns. Next we define and give some geometric properties of the L'-, L-, and λ-families of curves. The relation to wide-open trihorns is stated in the basic Theorems 7 and 11. Then we find the transformations which preserve these families and also show the non-existence of a transformation converting every L-family of curves into a λ-family of curves. In the final part of our paper we prove that M_{12} and μ_{12} uniquely characterize respectively the direct conformal group and the direct equi-long group within the group of all contact transformations, and that there is no contact transformation which changes every M_{12} into an equal μ_{12}. The relation to Finsler metric (or space) is stated at the end.

The author wishes to thank John De Cicco for valuable assistance in writing this paper.

1. **The extended lemniscate.** Let C be a unit circle and C' any given circle such that C and C' have a point O in common, and let k be a given number. Through O draw any line l. Then l will meet C and C' in Q and Q', respectively. On l, let us take the point R so that $\vec{O}R = k\vec{O}Q$. Finally on l, we take the points P_i $(i = 1, 2)$ so that $\overline{RP}_i$ is equal to the mean proportion between $\vec{O}Q$ and $\vec{O}Q'$. The totality of points P_1, P_2 constitutes an *extended lemniscate.*

We note that if C and C' are orthogonal and if $k = 0$, then our extended lemniscate is an ordinary lemniscate.

The equation of an extended lemniscate is

$$
(1)\qquad \begin{aligned} &[x^2 + y^2]^2 - 4k[x \cos \alpha + y \sin \alpha][x^2 + y^2] \\ &+4[x \cos \alpha + y \sin \alpha][x(k^2 \cos \alpha - r \cos \beta) + y(k^2 \sin \alpha - r \sin \beta)] = 0, \end{aligned}
$$

where O is taken as origin, and where $(1, \alpha)$ and (r, β) are the polar coördinates of the centers of C and C', respectively.

From this equation we observe that O is the node of the extended lemniscate and that the tangent to C at O is one of the tangent lines of the extended lemniscate at O. We call this the *principal tangent element* of our extended lemniscate.

The equation of any extended lemniscate may be put in the form

$$
(2)\qquad [x^2 + y^2]^2 + [ax + by][c(x^2 + y^2) + dx + ey] = 0,
$$

where O is at the origin and $ax + by = 0$ is the line of the principal tangent element.

If the tangent lines at the node of an extended lemniscate are orthogonal, then we shall say that our extended lemniscate is an *orthogonal extended lemniscate.*

2. **The extended cissoid.** Let C be a fixed circle and OA a fixed chord of C. At A let us draw the tangent line t to C. Let l be any line through O. Then it will meet C and t in the points Q and Q', respectively. On l let us take the point P so that $\vec{OP} = \vec{QQ'}$. The set of points P is called an *extended cissoid*.

We note that, if the chord OA is a diameter, we obtain an ordinary cissoid.

The equation of an extended cissoid is

$$[x^2 + y^2][x \cos (\alpha - \beta) - y \sin (\alpha - \beta)] = 2r[x \cos \beta + y \sin \beta]^2, \tag{3}$$

where O is taken as origin, $(r, \alpha + \beta)$ are the polar coördinates of the center of C and β is the inclination of the chord.

From this equation we observe that O is the cusp of the extended cissoid and its tangent line at O is the chord OA.

Thus any extended cissoid takes the form

$$(x^2 + y^2)(ax + by) = (cx + dy)^2, \tag{4}$$

where O is at the origin and $cx + dy = 0$ is the equation of the fixed chord OA.

3. **The conformal measure M_{12}.** In our previous work it was shown that any horn angle of first order contact has the unique conformal invariant

$$M_{12} = \frac{(a_2 - a_1)^2}{b_2 - b_1}, \tag{5}$$

where a_i and b_i are the curvature and the rate of variation per unit length of arc of the curve C_i of the horn angle at the vertex of the horn angle. We call it the *conformal measure* of the horn angle.

A *trihorn* is defined to be three curves C_1, C_2, C_3, which pass through a given point in a common direction. The fundamental triangular inequality is

$$M_{12}M_{23}M_{31}(M_{12} + M_{23} + M_{31}) \leqq 0.$$

If a trihorn (C_1, C_2, C_3) satisfies the condition

$$M_{12} + M_{23} + M_{31} = 0, \tag{6}$$

then it is said to be *wide-open*.

4. **The equi-long measure μ_{12}.** We have previously shown that any horn angle of first order contact has the unique equi-long invariant

$$\mu_{12} = \frac{(\alpha_2 - \alpha_1)^2}{\beta_2 - \beta_1}, \tag{7}$$

where α_i and β_i are the radius of curvature and the rate of variation of curvature per unit radian measure of the angle that the tangent line makes with a fixed line of the curve C_i of the horn angle at the vertex of the horn angle. We call it the *equi-long measure* of the horn angle.

If a trihorn (C_1, C_2, C_3) satisfies the condition

$$\mu_{12} + \mu_{23} + \mu_{31} = 0, \tag{8}$$

then we shall say that the trihorn is *dual-wide-open.*

5. **The L'-family of curves.** A three-parameter family of curves is called an *L'-family of curves* if its differential equation is of the form

$$y''' = fy''^2 + gy'' + h, \tag{9}$$

where f, g, h are arbitrary functions of x, y, y'.

The L'-families of curves include the trajectories of dynamics (Kasner, *Trajectories of dynamics,* Transactions of the American Mathematical Society, 1906); the plane sections of a surface (*Dynamical trajectories and the ∞^3 plane sections of a surface,* Proceedings of the National Academy of Sciences, 1931); and the curvature trajectories (*Dynamical trajectories and curvature trajectories,* Bulletin of the American Mathematical Society, 1934).

THEOREM 1. *For a three-parameter family of curves to be an L'-family of curves, it is necessary and sufficient that*

$$b = Aa^2 + Ba + C, \tag{10}$$

where A, B, C are arbitrary functions of x, y, y', and a and b are the curvature and the rate of variation of curvature per unit length of arc.

THEOREM 2. *For a three-parameter family of curves to be an L'-family of curves, it is necessary and sufficient that the locus of the second centers of curvature of the ∞^1 curves which contain a given lineal element, constructed at the point of the element, be a cubical parabola (related to the tangent and normal lines of the given lineal element as the y- and x-axes, respectively), which contains the point of the element.*

THEOREM 3. *For a three-parameter family of curves to be an L'-family of curves, it is necessary and sufficient that the locus of the foci of the osculating parabolas of the ∞^1 curves which contain a lineal element, constructed at the point of the element, be either a circle passing through the point of the element or an extended lemniscate with the given element as the principal tangent element of the extended lemniscate.*

6. **The L-family of curves.** A three-parameter family of curves is called an *L-family of curves* if its differential equation is of the form

$$y''' = \frac{3y'}{1 + y'^2} y''^2 + gy'' + h, \tag{11}$$

where g and h are arbitrary functions of x, y, y'. Thus every L-family of curves is also an L'-family of curves.

THEOREM 4. *For a three-parameter family of curves to be an L-family of curves, it is necessary and sufficient that*

$$Aa + Bb + C = 0, \tag{12}$$

where A, B, C are arbitrary functions of x, y, y', and a and b are the curvature and the rate of variation of curvature per unit length of arc.

THEOREM 5. *For a three-parameter family of curves to be an L-family of curves, it is necessary and sufficient that the locus of the second centers of curvature of the ∞^1 curves which contain a given lineal element, constructed at the point of the element, be a cubical parabola (related to the tangent and normal lines of the given lineal element as the y- and x-axes, respectively), which passes through the point of the element in a direction orthogonal to the direction of the element.*

THEOREM 6. *For a three-parameter family of curves to be an L-family of curves, it is necessary and sufficient that the locus of the foci of the osculating parabolas of the ∞^1 curves which contain a lineal element, constructed at the point of the element, be either a circle passing through the point of the element and orthogonal to the element or an orthogonal extended lemniscate with the given element as the principal tangent element of the orthogonal extended lemniscate.*

THEOREM 7. *For a three-parameter family of curves to be an L-family of curves, it is necessary and sufficient that every trihorn of the family be wide-open.*

This last theorem indicates the immediate importance of *L*-families in conformal geometry.

7. **The λ-family of curves.** A three-parameter family of curves is called a *λ-family of curves* if its differential equation, in Hessian coördinates, is of the form

$$v''' = fv'' + g, \tag{13}$$

where f and g are arbitrary functions of u, v, v'.

In Cartesian coördinates, the differential equation of a λ-family of curves is

$$y''' = hy''^3 + ky''^2, \tag{14}$$

where h and k are arbitrary functions of x, y, y'.

THEOREM 8. *For a three-parameter family of curves to be a λ-family of curves, it is necessary and sufficient that*

$$A\alpha + B\beta + C = 0, \tag{15}$$

where A, B, C are arbitrary functions of u, v, v', and α and β are the radius of curvature and the rate of variation of the radius of curvature per unit radian measure of the angle that the tangent line makes with a fixed line.

THEOREM 9. *For a three-parameter family of curves to be a λ-family of curves, it is necessary and sufficient that the locus of the second centers of curvature of the ∞^1 curves which contain a given lineal element, constructed at the point of the element, be a straight line.*

THEOREM 10. *For a three-parameter family of curves to be a λ-family of curves, it is necessary and sufficient that the locus of the foci of the osculating parabolas, of the ∞^1 curves which contain a lineal element, constructed at the point of the element, be an extended cissoid which contains and has its cusp at the given element.*

THEOREM 11. *For a three-parameter family of curves to be a λ-family of curves, it is necessary and sufficient that every trihorn of the family be dual-wide-open.*

This theorem shows the importance of the concept of λ-family in general equi-long geometry.

8. L'-families of curves into L'-families of curves.

THEOREM 12. *Any contact transformation which converts every L'-family of curves into an L'-family of curves must be a point transformation.*

Let

$$X = \phi(x, y, p), \qquad Y = \psi(x, y, p), \qquad P = \chi(x, y, p) \tag{16}$$

be a contact transformation, so that

$$\frac{\psi_x + p\psi_y}{\phi_x + p\phi_y} = \frac{\psi_p}{\phi_p} = \chi. \tag{17}$$

The extended form of this contact transformation is given by the equation

$$\begin{aligned} \frac{dP}{dX} &= \frac{\chi_x + p\chi_y + p'\chi_p}{\phi_x + p\phi_y + p'\phi_p}, \\ \frac{d^2P}{dX^2} &= \frac{A + Bp' + Cp'^2 + Dp'^3 + Ep''}{(\phi_x + p\phi_y + p'\phi_p)^3}, \end{aligned} \tag{18}$$

where

$$\begin{aligned} A &= (\phi_x + p\phi_y)(\chi_{xx} + 2p\chi_{xy} + p^2\chi_{yy}) - (\chi_x + p\chi_y)(\phi_{xx} + 2p\phi_{xy} + p^2\phi_{yy}), \\ B &= (\phi_x + p\phi_y)(\chi_y + 2\chi_{xp} + 2p\chi_{yp}) + \phi_p(\chi_{xx} + 2p\chi_{xy} + p^2\chi_{yy}) \\ &\quad - (\chi_x + p\chi_y)(\phi_y + 2\phi_{xp} + 2p\phi_{yp}) - \chi_p(\phi_{xx} + 2p\phi_{xy} + p^2\phi_{yy}), \\ C &= (\phi_x + p\phi_y)\chi_{pp} + \phi_p(\chi_y + 2\chi_{xp} + 2p\chi_{yp}) \\ &\quad - (\chi_x + p\chi_y)\phi_{pp} - \chi_p(\phi_y + 2\phi_{xp} + 2p\phi_{yp}), \\ D &= \phi_p\chi_{pp} - \chi_p\phi_{pp}, \\ E &= \chi_p(\phi_x + p\phi_y) - \phi_p(\chi_x + p\chi_y). \end{aligned} \tag{19}$$

If we use the formulas

$$a = p'(1 + p^2)^{-3/2}, \qquad b = [(1 + p^2)p'' - 3pp'^2](1 + p^2)^{-3}$$

and the corresponding formulas for $\bar{a}$ and $\bar{b}$, the curvature and the rate of variation of curvature per unit length of arc of the transformed curve, the equations (18) become

$$\begin{aligned} \bar{a} &= \frac{(\chi_x + p\chi_y) + q^{3/2}\chi_p a}{(1 + \chi^2)^{3/2}[(\phi_x + p\phi_y) + q^{3/2}\phi_p a]}, \\ \bar{b} &= \frac{\bar{A} + q^{3/2}\bar{B}a + [q^3\bar{C} + 3pq^2\bar{E}]a^2 + q^{9/2}\bar{D}a^3 + q^2\bar{E}b}{(1 + \chi^2)^3[(\phi_x + p\phi_y) + q^{3/2}\phi_p a]^3}, \end{aligned} \tag{20}$$

where

$$
\begin{aligned}
\bar{A} &= (1 + \chi^2)A - 3\chi(\chi_x + p\chi_y)^2(\phi_x + p\phi_y), \\
\bar{B} &= (1 + \chi^2)B - 6\chi\chi_p(\chi_x + p\chi_y)(\phi_x + p\phi_y) - 3\chi\phi_p(\chi_x + p\chi_y)^2, \\
\bar{C} &= (1 + \chi^2)C - 3\chi\chi_p^2(\phi_x + p\phi_y) - 6\chi\chi_p\phi_p(\chi_x + p\chi_y), \\
\bar{D} &= (1 + \chi^2)D - 3\chi\chi_p^2\phi_p, \\
\bar{E} &= (1 + \chi^2)E, \qquad q = 1 + p^2.
\end{aligned} \tag{21}
$$

For a fixed lineal element (x, y, p), (20) is a transformation from (a, b) to $(\bar{a}, \bar{b})$. Then for every L'-family of curves to be carried into an L'-family of curves, the transformation (20) must carry every equation of the form (10) into an equation of the form (10). This can happen when and only when

$$\phi_p = 0, \qquad \bar{D} = 0. \tag{22}$$

By equations (17), (19), (21), (22), we see that the required transformation must be a point transformation, and our theorem is proved.

9. L-families of curves into L-families of curves.

THEOREM 13. *Any contact transformation which changes every L-family of curves into an L-family of curves must be a conformal transformation.*

If our contact transformation carries every L-family of curves into an L-family of curves, the transformation (20) must carry every equation of the form (12) into an equation of the form (12). It therefore follows that (20) must be a linear transformation, and therefore we must have

$$\phi_p = 0, \qquad \bar{D} = 0, \qquad (1 + p^2)\bar{C} + 3p\bar{E} = 0. \tag{23}$$

By equations (17), (19), (21), (23), we observe that our contact transformation must be a point transformation, and also we must have

$$(1 + p^2)\bar{C} + 3p\bar{E} = 0. \tag{24}$$

Upon simplifying (24) by means of (19) and (21), we obtain the single condition

$$
\begin{aligned}
&\phi_x(\phi_x\phi_y + \psi_x\psi_y) + (2\phi_x\phi_y^2 + \phi_y\psi_x\psi_y - \phi_x^3 - \phi_x\psi_x^2 + \phi_x\psi_y^2)p \\
&\quad + (\phi_y\psi_y^2 + \phi_y^3 - 2\phi_x^2\phi_y - \phi_x\psi_x\psi_y - \phi_y\psi_x^2)p^2 - \phi_y(\phi_x\phi_y + \psi_x\psi_y)p^3 = 0.
\end{aligned} \tag{25}
$$

Since (25) is an identity, we then obtain the four conditions

$$
\begin{aligned}
\phi_x(\phi_x\phi_y + \psi_x\psi_y) &= 0, \\
2\phi_x\phi_y^2 + \phi_y\psi_x\psi_y - \phi_x^3 - \phi_x\psi_x^2 + \phi_x\psi_y^2 &= 0, \\
\phi_y\psi_y^2 + \phi_y^3 - 2\phi_x^2\phi_y - \phi_x\psi_x\psi_y - \phi_y\psi_x^2 &= 0, \\
\phi_y(\phi_x\phi_y + \psi_x\psi_y) &= 0.
\end{aligned} \tag{26}
$$

If $\phi_x = 0$, then $\phi_y \neq 0$, and $\psi_x \neq 0$. From (26) we obtain $\psi_y = 0$, and $\phi_y = \pm\psi_x$. Thus our transformation is conformal.

If $\phi_y = 0$, then $\phi_x \neq 0$, and $\psi_y \neq 0$. From (26) we obtain $\psi_x = 0$, and $\phi_x = \pm\psi_y$. Thus our transformation is conformal.

Now we can assume that $\phi_x \neq 0$ and $\phi_y \neq 0$. The equations (26) are then equivalent to the equations

$$\phi_x\phi_y + \psi_x\psi_y = 0, \qquad \phi_x^2 + \psi_x^2 = \phi_y^2 + \psi_y^2,$$

which are obviously the conditions for conformality. This completes the proof of the theorem.

10. λ-families of curves into λ-families of curves.

THEOREM 14. *Any contact transformation which transforms every λ-family of curves into a λ-family of curves must be a line transformation of the form*

$$\phi = F(u), \quad \psi = G(u, v). \tag{27}$$

It is seen that this theorem is not the exact dual of Theorem 13. This group of transformations contains one arbitrary function of one variable and one arbitrary function of two variables. A subgroup of this group of transformations is the set of equi-long transformations.

Let

$$U = \phi(u, v, w), \qquad V = \psi(u, v, w), \qquad W = \chi(u, v, w) \tag{28}$$

be a contact transformation, so that

$$\frac{\psi_u + w\psi_v}{\phi_u + w\phi_v} = \frac{\psi_w}{\phi_w} = \chi. \tag{29}$$

The extended form of this contact transformation is given by the equations

$$\begin{aligned} \frac{dW}{dU} &= \frac{\chi_u + w\chi_v + w'\chi_w}{\phi_u + w\phi_v + w'\phi_w}, \\ \frac{d^2W}{dU^2} &= \frac{A + Bw' + Cw'^2 + Dw'^3 + Ew''}{(\phi_u + w\phi_v + w'\phi_w)^3}, \end{aligned} \tag{30}$$

where A, B, C, D and E are given by formulas of the same form (19).

If we use the formulas

$$\bar{\alpha} = V + \frac{dW}{dU}, \qquad \bar{\beta} = W + \frac{d^2W}{dU^2}, \qquad \alpha = v + w', \qquad \beta = w + w'',$$

the equations (30) become

$$\begin{aligned} \bar{\alpha} &= \frac{(\chi_u + w\chi_v) + \psi(\phi_u + w\phi_v) - v(\chi_w + \psi\phi_w) + \alpha(\chi_w + \psi\phi_w)}{\phi_u + w\phi_v - v\phi_w + \alpha\phi_w}, \\ (\phi_u + w\phi_u - v\phi_w + \alpha\phi_w)^3\bar{\beta} &= (\bar{A} - v\bar{B} + v^2\bar{C} - v^3\bar{D} - Ew) \\ &\quad + \alpha(\bar{B} - 2v\bar{C} + 3v^2\bar{D}) + \alpha^2(\bar{C} - 3v\bar{D}) + \alpha^3\bar{D} + \beta E, \end{aligned} \tag{31}$$

where

$$
(32)\qquad
\begin{aligned}
\bar{A} &= A + \chi(\phi_u + w\phi_v)^3,\\
\bar{B} &= B + 3\chi(\phi_u + w\phi_v)^2\phi_w\,,\\
\bar{C} &= C + 3\chi(\phi_u + w\phi_v)\phi_w^2\,,\\
\bar{D} &= D + \chi\phi_w^3\,.
\end{aligned}
$$

For a fixed lineal element (u, v, w) the equations (31) define a correspondence from (α, β) to $(\bar{\alpha}, \bar{\beta})$. If our contact transformation carries every λ-family of curves into a λ-family of curves, the transformation (31) must carry every equation of the type (15) into an equation of the type (15). It therefore follows that (31) must be a linear transformation, and hence

$$
(33)\qquad \phi_w = 0, \qquad \bar{D} = 0, \qquad \bar{C} - 3v\bar{D} = 0.
$$

From equations (19), (29) and (32) we see that our contact transformation must be a line transformation, and also we must have

$$
(34)\qquad C = 0.
$$

Upon simplifying (34) by means of (19) and (29), we obtain the single condition

$$
(35)\qquad -3J\phi_v = 0,
$$

where J is the Jacobian of the transformation. Hence $\phi_v = 0$, from which we obtain immediately the equations (27). Theorem 14 is completely proved.

From Theorems 13 and 14 we obtain

THEOREM 15. *The group of contact transformations which change every L-family of curves into an L-family of curves and every λ-family of curves into a λ-family of curves is the group of rigid motions, reflections and magnifications, that is, the similitude group.*

11. Impossibility of converting L-families into λ-families.

THEOREM 16. *There is no contact transformation which changes every L-family of curves into a λ-family of curves.*

By means of the obvious relations

$$
\alpha = \frac{1}{a}, \qquad \beta = -\frac{b}{a^3},
$$

the equations (20) may be written in the form

$$
(36)\qquad
\begin{aligned}
\bar{\alpha} &= \frac{(1 + \chi^2)^{\frac{1}{2}}[(\phi_x + p\phi_y) + q^{\frac{3}{2}}\phi_p a]}{(\chi_x + p\chi_y) + q^{\frac{3}{2}}\chi_p a},\\
\bar{\beta} &= -\frac{(1 + \chi^2)^{\frac{3}{2}}[\bar{A} + q^{\frac{3}{2}}\bar{B}a + \{q^3\bar{C} + 3pq^2\bar{E}\}a^2 + q^{\frac{9}{2}}\bar{D}a^3 + q\bar{E}b]}{[(\chi_x + p\chi_y) + q^{\frac{3}{2}}\chi_p a]^3}.
\end{aligned}
$$

If our contact transformation carries every L-family of curves into a λ-family of curves, then (36) must carry every equation of type (12) into an equation of type (15). Thus (36) must be a linear transformation, and therefore

$$(37) \qquad \chi_p = 0, \qquad \bar{D} = 0, \qquad (1 + p^2)\bar{C} + 3p\bar{E} = 0.$$

From (17), (19), (21), (36), (37) we see that our contact transformation must be of the form

$$(38) \qquad \begin{aligned} \phi &= \frac{g_x + pg_y}{f_x + pf_y}, \\ \psi &= f\phi - g, \\ \chi &= f, \end{aligned}$$

where f and g are arbitrary functions of x and y only, and where

$$(39) \qquad (1 + p^2)[\phi_p\chi_y - (\chi_x + p\chi_y)\phi_{pp}] - 3p\phi_p(\chi_x + p\chi_y) = 0.$$

Substituting (38) into (39) and simplifying, we obtain

$$(40) \qquad (-pf_x + f_y)J = 0,$$

where $J = f_xg_y - f_yg_x$. Now $J \neq 0$, for otherwise by (38) ϕ, ψ, χ would be independent of p. Hence (40) becomes

$$(41) \qquad -pf_x + f_y = 0.$$

Since (41) is an identity, we obtain f = constant. By (38) this is impossible, and hence there is no contact transformation which converts every L-family of curves into a λ-family of curves.

12. **Conformal transformations and horn angles.** We now prove that the conformal measure M_{12} completely characterizes the group of direct conformal transformations among all contact transformations.

THEOREM 17. *If a contact transformation leaves invariant the conformal measure M_{12} of every horn angle, then it must be a direct conformal transformation.*

For then every wide-open trihorn must be carried into a wide-open trihorn, and hence by Theorem 7, every L-family of curves must be transformed into an L-family of curves. By Theorem 13, our transformation must be conformal. Thus equations (20) assume the form

$$(42) \qquad \begin{aligned} \bar{a} &= ma + h, \\ \bar{b} &= \pm m^2 b + k, \end{aligned}$$

where we take the plus or minus sign according as the correspondence is direct or reverse conformal. Then for M_{12} to be invariant we must take the plus sign, and hence the required transformation is direct conformal.

13. **Equi-long transformations and horn angles.** Finally we prove that the equi-long measure μ_{12} completely characterizes the group of direct equi-long transformations among all contact transformations.

THEOREM 18. *If a contact transformation leaves invariant the equi-long measure μ_{12} of every horn angle, then it must be a direct equi-long transformation.*

For then every dual-wide-open trihorn must be transformed into a dual-wide-open trihorn, and hence by Theorem 11 every λ-family of curves must be transformed into a λ-family of curves. By Theorem 14 our transformation must be a line transformation of the form (27). For such a transformation the equations (31) become

$$\bar{\alpha} = \frac{(\chi_u + w\chi_v) + \psi\phi_u - v\chi_w + \alpha\chi_w}{\phi_u}, \tag{43}$$
$$\bar{\beta} = \frac{(\bar{A} - v\bar{B} - w\chi_w\phi_u) + \alpha\bar{B} + \beta\chi_w\phi_u}{\phi_u^3},$$

where

$$\bar{A} = A + \chi\phi_u^3,$$
$$\bar{B} = B = \phi_u(\chi_v + 2\chi_{uw} + 2w\chi_{vw}) - \chi_w\phi_{uu}.$$

For (43) to leave μ_{12} invariant, we must have

$$\phi_u(\chi_v + 2\chi_{uw} + 2w\chi_{vw}) - \chi_w\phi_{uu} = 0, \tag{44}$$
$$\frac{\chi_w}{\phi_u^2} = \left(\frac{\chi_w}{\phi_u}\right)^2.$$

Since $\phi_u \neq 0$, and $\chi_w \neq 0$, the second of equations (44) shows that

$$\chi_w = 1. \tag{45}$$

From (27), (29) and (45) we obtain immediately that our contact transformation is a direct equi-long transformation. Of course it is obvious that our direct equi-long transformation satisfies the first of equations (44). Theorem 18 is completely proved.

From Theorems 17 and 18 we obtain

THEOREM 19. *If a contact transformation leaves every conformal measure M_{12} and every equi-long measure μ_{12} invariant, then it must be a rigid motion.*

14. **Not every M_{12} into a μ_{12}.**

THEOREM 20. *There is no contact transformation which carries every horn angle of conformal measure M_{12} into a horn angle with an equal equi-long measure μ_{12}.*

For then every wide-open trihorn must be carried into a dual-wide-open trihorn, and hence by Theorems 7 and 11, the transformation must carry every L-family of curves into a λ-family of curves. By Theorem 16 this is impossible.

15. **Finsler metric.** Since M and μ have the same algebraic form $(x_2 - x_1)^2/(y_2 - y_1)$, the conformal and equi-long theories of horn angles lead to the same abstract metric. *This may in fact be considered as a special Finsler metric (Finsler space), defined by the integral*

$$S = \int \frac{dx^2}{dy} = \int \frac{1}{y'}\, dx.$$

We thus obtain essentially the same trihornometry in the two theories.

Theorem 20 asserts that it is impossible to pass from the conformal plane (x, y) to the equi-long plane (u, v) by any contact transformation. The auxiliary planes (a, b) and (α, β) may be related, of course, by a transformation of differential elements of third order.

We remark in conclusion that relative invariants, for example Ostrowski's $\gamma_2 - \gamma_1$, or our expression of third order, $d\gamma_2/ds_2 - d\gamma_1/ds_1$, may serve to characterize completely the conformal group. The dual results are also valid.

Columbia University.

QUATERNARY CREMONA GROUPS OF TERNARY TYPE

By Frank C. Gentry

Introduction. We consider the possibility of using involutions determined by webs of quartic surfaces of degree 2 as generators of groups of Cremona transformations in space. Coble[1] has discussed the same problem using involutions determined by webs of cubic surfaces as generators.

For a web of quartic surfaces of degree 2 to contain in its base a curve of index numbers[2] (α_0', α_1') and of multiplicity i, a simple curve (α_0, α_1) meeting the multiple curve s times, B_j j-fold points $(j = 1, 2, 3, \cdots)$, Hudson[3] gives for the postulation P and the equivalence E the formulas:

$$P = \frac{i(i+1)}{12}\{36\alpha_0' + (2i+1)\alpha_1'\} + 6\alpha_0 + \frac{\alpha_1}{2} - is + \sum_j \frac{j(j+1)(j+2)B_j}{6} = 31,$$

$$E = i^2(12\alpha_0' + i\alpha_1') + 12\alpha_0 + \alpha_1 - (3i-1)s + \sum_j j^3 B_j = 62.$$

The following solutions of these equations, for $i > 1$, $\alpha_0' \neq 0$, $B_1 \neq 0$, lead to webs of non-degenerate quartic surfaces:

No.	i	α_0'	α_1'	α_0	α_1	s	B_1	B_2	No.	i	α_0'	α_1'	α_0	α_1	s	B_1	B_2
I	2	1	-2	6	-18	5	1	0	VIII	2	1	-2	2	-4	2	4	2
II	2	1	-2	7	-30	5	1	0	IX	2	1	-2	0	0	0	6	3
III	2	1	-2	5	-12	4	2	0	X	2	2	-6	4	-16	4	2	0
IV	2	1	-2	6	-24	4	2	0	XI	2	2	-6	3	-10	3	3	0
V	2	1	-2	5	-18	3	3	0	XII	2	2	-6	2	-4	2	4	0
VI	2	1	-2	4	-12	2	4	0	XIII	2	2	-6	0	0	0	6	1.
VII	2	1	-2	4	-8	4	2	1									

Sharpe and Snyder[4] have determined the homaloidal webs and fundamental and principal elements of the involutions of Cases II, IV, V and VI. Except in

Received October 18, 1937.

[1] A. B. Coble, *Groups of Cremona transformations in space of planar type*, I and II, this Journal, vol. 2 (1936), pp. 1, 205.

[2] A. B. Coble, *Restricted systems of equations*, I, II, American Journal of Mathematics, vol. 36 (1914), pp. 167, 295.

[3] Hilda P. Hudson, *Cremona Transformations in Plane and Space*, Chapter XI, Cambridge, 1927.

[4] F. R. Sharpe and V. Snyder, *Certain types of involutorial space transformations*, Transactions of the American Mathematical Society, vol. 21 (1920), pp. 52–78.

Case X, we shall use their methods for obtaining a description of the involutions of the other cases. Then, in each case, by allowing the simple base points to vary while the remainder of the base is held fixed, we shall generate groups of Cremona transformations isomorphic to certain linear groups.

I. *The involution I_1 determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_4\, \epsilon_1\, \epsilon_2\, p_1)^4$ where the double line ϵ_0 is trisecant to the rational quartic curve C_4 and across the two lines ϵ_1, ϵ_2.* Two members of the web W meet in ϵ_0^4, C_4, ϵ_1, ϵ_2 and an elliptic sextic curve C_6 on p_1 and meeting ϵ_0 3 times, C_4 9 times and ϵ_1, ϵ_2 each 3 times. The web W contains 3 pencils of degenerate members: $W_0 = (\epsilon_0\, C_4)^2(\epsilon_0\, \epsilon_1\, \epsilon_2\, p_1)^2$, $W_j = (\epsilon_0\, \epsilon_j)^1(\epsilon_0\, C_4\, \epsilon_h\, p_1)^3$ $(j, h = 1, 2; j \neq h)$. The plane $P_{K_j}(\epsilon_0\, \epsilon_j)^1$ is a P-surface of the F-curve of the first kind K_j, of order 2, on p_1, 4-secant to C_4 and 2-secant to ϵ_h. The quadric $P_L(\epsilon_0\, C_4)^2$ is a P-surface of the F-curve L, of the first kind, of order 1, on p_1 and across ϵ_1, ϵ_2, K_1 and K_2. The P-surfaces P_{K_1}, P_{K_2}, P_L and the surface of coincident points $R(\epsilon_0^4\, C_4^2\, \epsilon_1^2\, \epsilon_2^2\, p_1^2)^8$ make up the Jacobian of the web W. R meets P_{K_1}, P_{K_2} and P_L, respectively, in 2 lines ρ_1, 2 lines ρ_2 and 2 conics θ which are therefore self-corresponding F-curves of the second kind.

The homaloidal web H and the other P-surfaces of the involution are determined by making use of the (1, 2) transformation $y_i = Q_4^{(i)}(x)$ $(i = 1, 2, 3, 4)$, where $Q_4^{(i)}(x)$ are 4 linearly independent members of the web W. By considering the intersection of these surfaces with one another and with the surface R, we find that the involution I_1 possesses the additional self-corresponding F-curves of the second kind: the line r on p_1 and across ϵ_0 and C_4; the 6 lines s across ϵ_1, ϵ_2 and bisecant to C_4; the 4 conics C_2 on p_1 and meeting ϵ_0, ϵ_1, ϵ_2 each once and C_4 3 times; and the rational cubic C_3 on p_1 and meeting ϵ_0 once, ϵ_1, ϵ_2 each twice and C_4 5 times.

The characteristics of the homaloidal web H and the P-surfaces with respect to the base $(n\colon L, K_j, K_h, \epsilon_0, C_4, \epsilon_j, \epsilon_h, p_1; \rho_j, r, s, \theta, C_2, C_3)$ are as follows:

$$\begin{array}{ll}
H & (17\colon 2, 1, 1,\ 8, 4, 5, 5, 6; 1, 1, 1, 2, 2, 3),\\
P_L & (\ 2\colon 0, 0, 0,\ 1, 1, 0, 0, 0; 0, 0, 0, 1, 0, 0),\\
P_{K_j} & (\ 1\colon 0, 0, 0,\ 1, 0, 1, 0, 0; 1, 0, 0, 0, 0, 0),\\
P_{\epsilon_0} & (\ 8\colon 1, 1, 1,\ 4, 2, 2, 2, 3; 1, 1, 0, 1, 1, 1),\\
P_{C_4} & (26\colon 4, 1, 1, 12, 6, 8, 8, 9; 1, 1, 2, 4, 3, 5),\\
P_{\epsilon_j} & (\ 9\colon 1, 1, 0,\ 4, 2, 3, 3, 3; 1, 0, 1, 1, 1, 2),\\
P_{p_1} & (\ 4\colon 0, 0, 0,\ 2, 1, 1, 1, 2; 0, 1, 0, 0, 1, 1),
\end{array}$$

where $j, h = 1, 2$ $(j \neq h)$.

Let a surface s_z have the characteristic $(z\colon \bar{z}_0, \bar{z}_1, \bar{z}_2, z_0, z_*, z_1, z_2, x_1)$ with respect to the involution I_1, where z is the order of s_z and $\bar{z}_0$, $\bar{z}_1$, $\bar{z}_2$, z_0, z_*, z_1, z_2, and x_1 are its multiplicities on the F-curves $L, K_1, K_2, \epsilon_0, C_4, \epsilon_1, \epsilon_2$ and at the F-point p_1, respectively. The surface s_z is transformed by the involution I_1 into a surface $s'_{z'}$ whose characteristic is given in terms of the original characteristic by means of a linear transformation obtained immediately from the description of I_1 given above.

Let G be the group of Cremona transformations generated by involutions I_a with fixed F-curves ϵ_0, C_4, ϵ_1, ϵ_2 and variable F-point p_a chosen from the ρ points p_i $(i = 1, 2, \cdots, \rho)$, and hence the variable F-curves L, K_1, K_2. Let the points p_i be generically placed with respect to the fixed F-curves and the fixed P-surfaces of the variable F-curves. In forming products of such generators we suppose that the F-point p_a of the last factor may fall on one of the variable F-points of the preceding product but that, aside from such incidence, it is in generic position with respect to such F-points. The additional base points of these involutions would appear in the characteristic of s_z and in the transformation giving the characteristic of $s'_{z'}$ we should have the additional equations: $x'_i = x_i$ $(i > 1)$.

If in the linear transformation on the characteristic of s_z we make the substitutions: $u_0 = 2z - z_0 - 4z_* - z_1 - z_2$, $u_1 = z - z_0 - z_* - z_1$, $u_2 = z - z_0 - z_* - z_2$, $v = z - 2z_* - z_1 - z_2$, $t_0 = 6z - 3z_0 - 9z_* - 3z_1 - 3z_2$, $t_i = x_i$; then aside from $u'_j = \bar{z}_j$, $u_j = \bar{z}'_j$ $(j = 0, 1, 2)$ and $v' = -v$, the transformation becomes the involution

$$i_1: \quad \begin{aligned} t'_0 &= 2t_0 - 3t_1, \\ t'_1 &= t_0 - 2t_1, \\ t'_i &= t_i \qquad (i > 1). \end{aligned}$$

Let $g_\rho(2)$ be the group of linear transformations generated by involutions i_a $(a = 1, 2, 3, \cdots, \rho)$. Evidently $g_\rho(2)$ belongs to the type $g_\rho(\alpha)$ for $r = 1$ discussed by Coble.[5]

If the generic element of the linear group $g_\rho(2)$ is

$$g': \quad \begin{aligned} t'_0 &= \alpha_{00}t_0 - \alpha_{01}t_1 - \cdots - \alpha_{0\rho}t_\rho, \\ t'_i &= \alpha_{i0}t_0 - \alpha_{i1}t_1 - \cdots - \alpha_{i\rho}t_\rho \qquad (i = 1, 2, \cdots, \rho), \\ t'_j &= t_j \qquad (j > \rho); \end{aligned}$$

then by a comparison of particular products in the groups G and $g_\rho(2)$ we infer that the corresponding element G' of the Cremona group G has a homaloidal web H and P-surfaces whose characteristics with respect to the base (n: L, K_1, K_2, ϵ_0, C_4, ϵ_1, ϵ_2, p_1, p_2, $\cdots$, p_ρ) are:

$$\begin{array}{ll}
H & (\ 8\alpha_{00} \pm 1: 2, 1, 1, 4\alpha_{00}, 2\alpha_{00}, 2\alpha_{00} \pm 1, 2\alpha_{00} \pm 1, 6\alpha_{10}, 6\alpha_{20}, \cdots, 6\alpha_{\rho0}), \\
P_{\epsilon_0} & (\ 4\alpha_{00} \quad : 1, 1, 1, 2\alpha_{00}, \alpha_{00}, \alpha_{00}, \alpha_{00}, 3\alpha_{10}, 3\alpha_{20}, \cdots, 3\alpha_{\rho0}), \\
P_{C_4} & (12\alpha_{00} \pm 2: 4, 1, 1, 6\alpha_{00}, 3\alpha_{00}, 3\alpha_{00} \pm 2, 3\alpha_{00} \pm 2, 9\alpha_{10}, 9\alpha_{20}, \cdots, 9\alpha_{\rho0}), \\
P_{\epsilon_1} & (\ 4\alpha_{00} \pm 1: 1, 1, 0, 2\alpha_{00}, \alpha_{00}, \alpha_{00} \pm 1, \alpha_{00} \pm 1, 3\alpha_{10}, 3\alpha_{20}, \cdots, 3\alpha_{\rho0}), \\
P_{\epsilon_2} & (\ 4\alpha_{00} \pm 1: 1, 0, 1, 2\alpha_{00}, \alpha_{00}, \alpha_{00} \pm 1, \alpha_{00} \pm 1, 3\alpha_{10}, 3\alpha_{20}, \cdots, 3\alpha_{\rho0}), \\
P_{p_i} & (\ 4k \quad : 0, 0, 0, 2k, k, k, k, \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i}),
\end{array}$$

[5] A. B. Coble, *A class of linear groups with integral coefficients*, this Journal, vol. 3 (1937), pp. 175–199.

$(i = 1, \cdots, \rho)$; where $\alpha_{0i} = 3k$, k an integer, and where L, K_1, K_2 are variable F-curves of the first kind corresponding to the fixed P-surfaces $P_L(\epsilon_0\, C_4)^2$, $P_{K_1}(\epsilon_0\, \epsilon_1)^1$, $P_{K_2}(\epsilon_0\, \epsilon_2)^1$, respectively.

The theorem is true for i_1 and I_1. Hence in order to complete the proof we need only show that $g'i_1 \sim G'I_1$ and $g'i_{\rho+1} \sim G'I_{\rho+1}$. This follows immediately from a comparison of the products indicated.

II. *The involution I_1 determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_7\, p_1)^4$ where the double line ϵ_0 is 5-secant to the septimic curve of genus 2 C_7.* Two members of the web W meet in ϵ_0^4, C_7 and an elliptic quintic curve C_5 on p_1 and meeting ϵ_0 3 times and C_7 11 times. W contains the uniquely determined degenerate member $W_1 = (\epsilon_0\, p_1)^1(\epsilon_0\, C_7)^3$ whose factors are interchanged by the involution I_1. The quartic surface $P_K(\epsilon_0^3\, C_7)^4$ is a P-surface of an F-curve K, of the first kind, of order 5, 3-fold at p_1 and meeting C_7 14 times. The surface of coincident points is $R(\epsilon_0^4\, C_7^2\, p_1^2)^8$. The involution possesses the following self-corresponding F-curves of the second kind: the 6 lines ρ meeting C_7 twice and ϵ_0 once and on R and P_K, the 2 lines r on p_1 and across C_7 and ϵ_0, the 4 lines t quadrisecant to C_7 and the 4 conics C_2 on p_1 and meeting ϵ_0 once and C_7 5 times. The characteristics of the homaloidal web H and the P-surfaces with respect to the base $(n\colon K, \epsilon_0, C_7, p_1;\ \rho, r, t, C_2)$ are as follows:

$$H(15\colon 1, 7, 4, 5; 1, 1, 1, 2),$$

$$P_K(4\colon 0, 3, 1, 0; 1, 0, 0, 0), \qquad P_{\epsilon_0}(8\colon 1, 4, 2, 3; 1, 1, 0, 1),$$

$$P_{C_7}(36\colon 2, 16, 10, 11; 2, 1, 4, 5), \qquad P_{p_1}(4\colon 0, 2, 1, 2; 0, 1, 0, 1).$$

Let G be the group of Cremona transformations generated by involutions I_a with fixed F-curves ϵ_0, C_7 and variable F-point p_a chosen from the ρ points p_i $(i = 1, 2, 3, \cdots, \rho)$ and hence the variable F-curve K. We restrict the points p_i as in Case I.

The surface s_z having the characteristic $(z; \bar{z}, z_0, z_1, x_1)$, where z is the order of s_z and $\bar{z}, z_0, z_1, x_1$ are its multiplicities on the F-curves K, ϵ_0, C_7 and at the F-point p_1, is carried by I_1 into the surface $s'_{z'}$ whose characteristic is given in terms of the original characteristic by a linear transformation. If in this transformation we make the substitution $u_1 = z - z_0 - 2z_1$, $u_2 = z - 4z_1$, $t_0 = 5z - 3z_0 - 11z_1$, $t_i = x_i$; then aside from $u_1' = \bar{z}$, $u_1 = \bar{z}'$ and $u_2' = -u_2$, it becomes the transformation i_1 of Case I and the linear group generated by involutions i_a $(a = 1, 2, 3, \cdots, \rho)$ is a $g_\rho(2)$. If the generic element g' of $g_\rho(2)$ is defined as in Case I, then the homaloidal web H and the P-surfaces of the corresponding element G' of the Cremona group G have the following characteristics with respect to the base $(n\colon K, \epsilon_0, C_7, p_1, p_2, \cdots, p_\rho)$:

$$\begin{array}{llllllll}
H\ (20k \pm 5\colon & 1, & 10k \pm 3, & 5k \pm 1, & 5\alpha_{10}, & 5\alpha_{20}, & \cdots, & 5\alpha_{\rho 0}),\\
P_{\epsilon_0}(\ 4\alpha_{00} & \colon 1, & 2\alpha_{00} & , \alpha_{00} & , 3\alpha_{10}, & 3\alpha_{20}, & \cdots, & 3\alpha_{\rho 0}),\\
P_{C_7}(44k \pm 8\colon & 2, & 22k \pm 6, & 11k \pm 1, & 11\alpha_{10}, & 11\alpha_{20}, & \cdots, & 11\alpha_{\rho 0}),\\
P_{p_i}(\ 4j & \colon 0, & 2j & , j & , \alpha_{1i}, & \alpha_{2i}, & \cdots, & \alpha_{\rho i}),
\end{array}$$

($i = 1, 2, \cdots, \rho$), where $\alpha_{00} = 3k \pm 1$, $\alpha_{0i} = 3j$, k and j integers; and where K is a variable F-curve of the first kind corresponding to the fixed P-surface $P_K(\epsilon_0^3\, C_7)^4$. The proof is the same as that for the theorem of Case I.

III. *The involution I_{12} determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_2\, \epsilon_1\, \epsilon_2\, \epsilon_3\, p_1\, p_2)^4$ where the double line ϵ_0 meets the conic C_2 and each of the lines ϵ_1, ϵ_2, ϵ_3 once.* Two members of the web W meet in ϵ_0^4, C_2, ϵ_1, ϵ_2, ϵ_3 and an elliptic septimic curve C_7 on p_1 and p_2 and meeting ϵ_0 4 times, C_2 7 times and ϵ_j ($j = 1, 2, 3$) 3 times. W contains the three pencils of degenerate members $W_j = (\epsilon_0\, \epsilon_j)^1(\epsilon_0\, C_2\, \epsilon_m\, \epsilon_n\, p_1\, p_2)^3$ ($j, m, n = 1, 2, 3; j \neq m \neq n$); the uniquely determined degenerate member $(\epsilon_0\, C_2\, p_1\, p_2)^2(\epsilon_0\, \epsilon_1\, \epsilon_2\, \epsilon_3)^2$; and the pencil of degenerate members $(C_2)^1(\epsilon_0^2\, \epsilon_1\, \epsilon_2\, \epsilon_3\, p_1\, p_2)^3$. The planes $P_L(C_2)^1$, $P_{K_j}(\epsilon_0\, \epsilon_j)^1$ ($j = 1, 2, 3$) are P-surfaces corresponding to the F-curves of the first kind L, K_j. L is of order 2, is on p_1 and p_2, and meets ϵ_0, ϵ_1, ϵ_2, ϵ_3 each once. K_j is of order 3, is on p_1 and p_2 and meets ϵ_0 once, C_2 3 times and ϵ_m, ϵ_n each twice. The surface of coincident points is $R(\epsilon_0^4\, C_2^2\, \epsilon_1^2\, \epsilon_2^2\, \epsilon_3^2\, p_1^2\, p_2^2)^8$. The involution I_{12} possesses the following F-curves of the second kind: the 3 sets of 2 lines ρ_j ($j = 1, 2, 3$) across C_2, ϵ_0, ϵ_j and on R and P_{K_j}; the 3 lines r across C_2, ϵ_1, ϵ_2 and ϵ_3; the two lines s_i ($i = 1, 2$) on p_i and across ϵ_0 and C_2; the 2 conics θ meeting C_2 3 times and ϵ_0, ϵ_1, ϵ_2, ϵ_3 each once and on R and P_L; the 6 conics C_{ik} ($i = 1, 2; k = 1, 2, 3$) on p_i, and meeting ϵ_0, ϵ_1, ϵ_2, ϵ_3 each once and C_2 twice; the rational cubic C_3 on p_1, p_2 and meeting ϵ_0 twice, C_2 3 times, and ϵ_1, ϵ_2, ϵ_3 each once; and the rational quartic C_4 on p_1, p_2 and meeting ϵ_0, ϵ_1, ϵ_2, ϵ_3 each twice and C_2 4 times. The characteristics of the homaloidal web H and the P-surfaces with respect to the base (n: L, K_j, K_m, K_n, ϵ_0, C_2, ϵ_j, ϵ_m, ϵ_n, p_i, p_h; ρ_j, r, s_i, θ, C_{ik}, C_3, C_4)($i, h = 1, 2; i \neq h; j, m, n = 1, 2, 3; j \neq m \neq n$), are as follows:

$$
\begin{array}{ll}
H & (22\colon 2, 1, 1, 1, 11, 5, 6, 6, 6, 7, 7; 1, 1, 1, 2, 2, 3, 4),\\
P_L & (\ 1\colon 0, 0, 0, 0, \ 0, 1, 0, 0, 0, 0, 0; 0, 0, 0, 1, 0, 0, 0),\\
P_{K_j} & (\ 1\colon 0, 0, 0, 0, \ 1, 0, 1, 0, 0, 0, 0; 1, 0, 0, 0, 0, 0, 0),\\
P_{\epsilon_0} & (12\colon 1, 1, 1, 1, \ 6, 3, 3, 3, 3, 4, 4; 1, 0, 1, 1, 1, 2, 2),\\
P_{C_2} & (22\colon 3, 1, 1, 1, 11, 5, 6, 6, 6, 7, 7; 1, 1, 1, 3, 2, 3, 4),\\
P_{\epsilon_j} & (10\colon 1, 1, 0, 0, \ 5, 2, 3, 3, 3, 3, 3; 1, 1, 0, 1, 1, 1, 2),\\
P_{p_i} & (\ 4\colon 0, 0, 0, 0, \ 2, 1, 1, 1, 1, 2, 1; 0, 0, 1, 0, 1, 1, 1).
\end{array}
$$

Let G be the group of Cremona transformations generated by involutions I_{ab} with fixed F-curves ϵ_0, C_2, ϵ_1, ϵ_2, ϵ_3 and variable F-points p_a, p_b selected from the p generic points p_i ($i = 1, 2, 3, \cdots, \rho$), and hence the variable F-curves L, K_1, K_2, K_3.

The surface s_z having the characteristic (z: $\bar{z}_0$, $\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$, z_0, z_*, z_1, z_2, z_3, x_1, x_2), where z is the order of s_z and $\bar{z}_0$, $\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$, z_0, z_*, z_1, z_2, z_3, x_1, x_2 are its multiplicities on the F-curves L, K_1, K_2, K_3, ϵ_0, C_2, ϵ_1, ϵ_2, ϵ_3 and at the F-points p_1, p_2, is carried by the involution I_{12} into a surface $s'_{z'}$, whose characteristic is given in terms of the original characteristic by means of a linear

transformation. If in this transformation we make the substitutions: $u_0 = 2z - z_0 - 3z_* - z_1 - z_2 - z_3$, $u_j = z - z_0 - z_* - z_j$ $(j = 1, 2, 3)$, $v = z - z_* - z_1 - z_2 - z_3$, $t_0 = 7z - 4z_0 - 7z_* - 3z_1 - 3z_2 - 3z_3$, $t_i = x_i$; then, aside from $u_0' = \bar{z}_0$, $u_0 = \bar{z}_0'$, $u_j' = \bar{z}_j$, $u_j = \bar{z}_j'$ and $v' = -v$, it becomes the involution

$$i_{12}: \quad \begin{aligned} t_0' &= 3t_0 - 4t_1 - 4t_2, \\ t_1' &= t_0 - 2t_1 - t_2, \\ t_2' &= t_0 - t_1 - 2t_2, \\ t_i' &= t_i \qquad (i > 2). \end{aligned}$$

Let $g_\rho(2, 1, 1)$ be the group generated by involutions i_{ab} $(a, b = 1, 2, 3, \cdots, \rho;$ $a \neq b)$. This group is readily seen to belong to the type of linear groups $g_\rho(r, \epsilon, e)$ discussed by Coble.[6] If the generic element g' of $g_\rho(2, 1, 1)$ be defined as was that of $g_\rho(\alpha)$ in Case I, then the homaloidal web H and the P-surfaces of the corresponding element G' of the Cremona group G have the following characteristics with respect to the base $(n\colon L, K_j, K_m, K_n, \epsilon_0, C_2, \epsilon_j, \epsilon_m, \epsilon_n, p_1, \cdots, p_\rho)$:

$$\begin{array}{lllllllllll} H & (7\alpha_{00} \mp 1: 2, 1, 1, 1, & 14g \pm 3, & 7g \pm 2, & 7g \pm 1, & 7g \pm 1, & 7g \pm 1, & 7\alpha_{10}, \cdots, 7\alpha_{\rho 0}), \\ P_{\epsilon_0} & (4\alpha_{00} \quad : 1, 1, 1, 1, & 2\alpha_{00}, & \alpha_{00}, & \alpha_{00}, & \alpha_{00}, & \alpha_{00}, & 4\alpha_{10}, \cdots, 4\alpha_{\rho 0}), \\ P_{C_2} & (7\alpha_{00} \mp 1: 3, 1, 1, 1, & 14g \pm 3, & 7g \pm 2, & 7g \pm 1, & 7g \pm 1, & 7g \pm 1, & 7\alpha_{10}, \cdots, 7\alpha_{\rho 0}), \\ P_{\epsilon_j} & (3\alpha_{00} \mp 1: 1, 1, 0, 0, & 6g \pm 1, & 3g \pm 1, & 3g, & 3g, & 3g, & 3\alpha_{10}, \cdots, 3\alpha_{\rho 0}), \\ P_{p_i} & (\ \alpha_{0i} \quad : 0, 0, 0, 0, & 2h, & h, & h, & h, & h, & \alpha_{1i}, \cdots, \alpha_{\rho i}), \end{array}$$

$(j, m, n = 1, 2, 3; j \neq m \neq n; i = 1, 2, 3, \cdots, \rho)$, where $\alpha_{00} = 4g \pm 1$, $\alpha_{0i} = 4h$, g and h integers, and where L, K_1, K_2, K_3 are variable F-curves corresponding to fixed P-surfaces. In order to prove this theorem we have only to show that $g'i_{12} \sim G'I_{12}$, $g'i_{1,\rho+1} \sim G'I_{1,\rho+1}$, and $g'i_{\rho+1,\rho+2} \sim G'I_{\rho+1,\rho+2}$. A comparison of the products shows this to be true.

IV. *The involution I_{12} determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_6\, p_1\, p_2)^4$ where the double line ϵ_0 is quadrisecant to the elliptic sextic curve C_6.* Two members of the web W meet in ϵ_0^4, C_6 and an elliptic sextic curve on p_1, p_2 and meeting ϵ_0 4 times and C_6 12 times. The web W contains 2 uniquely determined degenerate members $W_i = (\epsilon_0\, p_i)^1(\epsilon_0\, C_6\, p_j)^3$ $(i, j = 1, 2; i \neq j)$ whose factors are interchanged by the involution.

The ruled quartic surface $P_K = (\epsilon_0^3\, C_6)^4$ is a P-surface corresponding to the F-curve K, of the first kind, of order 11, 4-fold at p_1 and p_2, and meeting ϵ_0 4 times and C_6 24 times. The surface of coincident points is $R(\epsilon_0^4\, C_6^2\, p_1^2\, p_2^2)^8$. The involution I_{12} possesses the following self-corresponding F-curves of the second kind: the 8 lines ρ across ϵ_0 and bisecant to C_6 and on P_K and R; the

[6] Ibid.

4 lines r_{ij} on p_i and across ϵ_0 and C_6; the 2 lines s quadrisecant to C_6; the 4 conics C_{ij} on p_i and meeting ϵ_0 once and C_6 5 times; and the 2 rational cubics C_3 on p_1, p_2 and meeting ϵ_0 twice and C_6 6 times. The homaloidal web H and the P-surfaces of the involution I_{12} have characteristics with respect to the base $(n: K, \epsilon_0, C_6, p_i, p_j; \rho, r_{ij}, s, C_{ij}, C_3)$ as follows:

$$\begin{array}{l} H\ (19:\ 1,\ \ 9,\ \ 5,\ \ 6,\ \ 6;\ 1,1,1,2,3),\\ P_K(\ 4:\ 0,\ \ 3,\ \ 1,\ \ 0,\ \ 0;\ 1,0,0,0,0),\\ P_{\epsilon_0}(12:\ 1,\ \ 6,\ \ 3,\ \ 4,\ \ 4;\ 1,1,0,1,2),\\ P_{C_6}(40:\ 2,\ 18,\ 11,\ 12,\ 12;\ 2,1,4,5,6),\\ P_{p_i}(\ 4:\ 0,\ \ 2,\ \ 1,\ \ 2,\ \ 1;\ 0,1,0,1,1) \qquad (i,j=1,2). \end{array}$$

Let G be the group of Cremona transformations generated by the involutions I_{ab} with fixed F-curves ϵ_0, C_6 and variable F-points p_a, p_b selected from the ρ generic points p_i $(i = 1, \cdots, \rho)$, and hence the variable F-curve K.

A surface s_z having the characteristic $(z: \bar{z}, z_0, z_1, x_1, x_2)$, where z is the order of s_z and $\bar{z}, z_0, z_1, x_1$ and x_2 are its multiplicities on the F-curves K, ϵ_0, C_6 and at the F-points p_1 and p_2, respectively, is carried by the involution I_{12} into the surface $s'_{z'}$ whose characteristic is given in terms of the original characteristic by a linear transformation. If in this transformation we make the substitutions: $u = z - z_0 - 2z_1$, $t_0 = 6z - 4z_0 - 12z_1$; $t_i = x_i$; then, aside from $u' = \bar{z}$, $u = \bar{z}'$, the transformation becomes the linear transformation i_{12} of Case III. Evidently the group generated by involutions i_{ab} $(a, b = 1, 2, 3, \cdots, \rho;\ a \neq b)$ is the group $g_\rho(2, 1, 1)$ of Case III. Let g' be the generic element of $g_\rho(2, 1, 1)$. Then the homaloidal web H and the P-surfaces of the corresponding element G' of the Cremona group G have the following characteristics with respect to the base $(n: K, \epsilon_0, C_6, p_1, p_2, \cdots, p_\rho)$:

$$\begin{array}{llllll} H\ (12k \pm 5: & 1, 6k \pm 3, & 3k \pm 1, & 6\alpha_{10}, & 6\alpha_{20}, \cdots, & 6\alpha_{\rho 0}),\\ P_{\epsilon_0}(\ 4\alpha_{00} : & 1, 2\alpha_{00}, & \alpha_{00}, & 4\alpha_{10}, & 4\alpha_{20}, \cdots, & 4\alpha_{\rho 0}),\\ P_{C_6}(12\alpha_{00} \mp 4: & 2, 6\alpha_{00}, & 3\alpha_{00} \mp 2, & 12\alpha_{10}, & 12\alpha_{20}, \cdots, & 12\alpha_{\rho 0}),\\ P_{p_i}(\ 4h : & 0, 2h, & h, & \alpha_{1i}, & \alpha_{2i}, \cdots, & \alpha_{\rho i}) \end{array}$$

$(i = 1, 2, \cdots, \rho)$, where $\alpha_{00} = 2k \pm 1$, $\alpha_{0i} = 4h$, k and h integers, and where K is a variable F-curve of the first kind corresponding to the fixed P-surface $P_K = (\epsilon_0^3\, C_6)^4$. The proof is the same as for the theorem of Case III.

V. *The involution I_{123} determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_5\, p_1\, p_2\, p_3)^4$ where the double line ϵ_0 is trisecant to the rational quintic curve C_5.* Two members of the web W meet in ϵ_0^4, C_5 and an elliptic septimic curve C_7 on p_1, p_2, p_3 and meeting ϵ_0 5 times and C_5 13 times. W contains 3 uniquely determined degenerate members $W_i = (\epsilon_0\, p_i)^1\, (\epsilon_0\, C_5\, p_j\, p_k)^3$ $(i, j, k = 1, 2, 3;\ i \neq j \neq k)$ whose factors are interchanged by the involution.

The ruled quartic surface $P_K = (\epsilon_0^3\, C_5)^4$ is the P-surface of an F-curve K, of the first kind, of order 19, 5-fold at p_1, p_2 and p_3 and meeting ϵ_0 10 times and C_5 36 times. The surface of coincident points is $R(\epsilon_0^4\, C_5^2\, p_1^2\, p_2^2\, p_3^2)^8$. The involution possesses the following self-corresponding F-curves of the second kind: the 10 lines ρ on R and P_K across ϵ_0 and bisecant to C_5; the 3 sets of 2 lines each r_i on p_i and across ϵ_0 and C_5; the line s quadrisecant to C_5; the 3 conics C_i on p_i and meeting ϵ_0 once and C_5 5 times; and 3 rational cubics $C_3^{(i)}$ on p_j and p_k and meeting ϵ_0 twice and C_5 6 times; and the rational quartic C_4 on p_1, p_2, p_3 and meeting ϵ_0 3 times and C_5 7 times.

The homaloidal web H and P-surfaces of the involution I_{123} have characteristics with respect to the base $(n\colon K, \epsilon_0, C_5, p_i, p_j, p_k; \rho, r_i, s, C_i, C_3^{(i)}, C_3^{(j)}, C_3^{(k)}, C_4)$ as follows:

$$\begin{array}{l}
H\ (23\colon 1, 11, 6, 7, 7, 7; 1, 1, 1, 2, 3, 3, 3, 4),\\
P_K(\ 4\colon 0, 3, 1, 0, 0, 0; 1, 0, 0, 0, 0, 0, 0, 0),\\
P_{\epsilon_0}(16\colon 1, 8, 4, 5, 5, 5; 1, 1, 0, 1, 2, 2, 2, 3),\\
P_{C_5}(44\colon 2, 20, 12, 13, 13, 13; 2, 1, 4, 5, 6, 6, 6, 7),\\
P_{p_i}(\ 4\colon 0, 2, 1, 2, 1, 1; 0, 1, 0, 1, 0, 1, 1, 1).
\end{array}$$

Let G be the group of Cremona transformations generated by involutions I_{jkh} with fixed F-curves ϵ_0, C_5 and variable F-points p_j, p_k, p_h chosen from the ρ generic points p_i $(i = 1, 2, 3, \cdots, \rho)$, and hence the variable F-curve K.

A surface s_z with characteristic $(z\colon \bar{z}, z_0, z_1, x_1, x_2, x_3)$ with respect to the involution is transformed into a surface $s_{z'}$ whose characteristic is given in terms of the original characteristic by a linear transformation. If in this transformation we make the substitutions: $u = z - z_0 - 2z_1$, $t_0 = 7z - 5z_0 - 13z_1$, $t_i = x_i$; then, apart from $u' = \bar{z}$, $u = \bar{z}'$, it becomes the involution

$$i_{123}\colon\quad \begin{aligned} t_0' &= 4t_0 - 5t_1 - 5t_2 - 5t_3,\\ t_i' &= t_0 - t_1 - t_2 - t_3 - t_i && (i = 1, 2, 3),\\ t_j' &= t_j && (j > 3). \end{aligned}$$

Let $g_\rho(3, 1, 1)$ be the linear group generated by involutions i_{jkh} $(j, k, h = 1, 2, 3, \cdots, \rho; j \neq k \neq h)$. $g_\rho(3, 1, 1)$ belongs to the type $g_\rho(r, \epsilon, e)$ mentioned above. The generic element G' of the Cremona group G has the following characteristics with respect to the base $(n\colon K, \epsilon_0, C_5, p_1, p_2, \cdots, p_\rho)$ and in terms of the coefficients of the corresponding element of the linear group $g_\rho(3, 1, 1)$:

$$\begin{array}{l}
H\ (28f \pm 5\colon 1, 14f \pm 3, 7f \pm 1, 7\alpha_{10}, 7\alpha_{20}, \cdots, 7\alpha_{\rho 0}),\\
P_{\epsilon_0}(\ 4\alpha_{00}\colon 1, 2\alpha_{00}, \alpha_{00}, 5\alpha_{10}, 5\alpha_{20}, \cdots, 5\alpha_{\rho 0}),\\
P_{C_5}(52f \pm 8\colon 2, 26f \pm 6, 13f \pm 1, 13\alpha_{10}, 13\alpha_{20}, \cdots, 13\alpha_{\rho 0}),\\
P_{p_i}(\ 4g\colon 0, 2g, g, \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i})
\end{array}$$

$(i = 1, 2, \cdots, \rho)$, where $\alpha_{00} = 5f \pm 1$, $\alpha_{0i} = 5g$, f and g integers, and where K is a variable F-curve of the first kind corresponding to the fixed P-surface $P_K(\epsilon_0^3\, C_5)^4$.

To prove the theorem we show that $g'i_{123} \sim G'I_{123}$, $g'i_{12,\rho+1} \sim G'I_{12,\rho+1}$, $g'i_{1,\rho+1,\rho+2} \sim G'I_{1,\rho+1,\rho+2}$ and $g'i_{\rho+1,\rho+2,\rho+3} \sim G'I_{\rho+1,\rho+2,\rho+3}$.

VI. *The involution I_{1234} determined by the web of quartic surfaces $W = (\epsilon_0^2\, C_2\, C_2'\, p_1\, p_2\, p_3\, p_4)^4$ where each of the conics C_2, C_2' meets the double line ϵ_0 once.* Two members of the web W meet in ϵ_0^4, C_2, C_2' and an elliptic octavic curve C_8 on p_i $(i = 1, 2, 3, 4)$ and meeting ϵ_0 6 times and C_2, C_2' each 7 times. The web W contains the 12 uniquely determined degenerate members: $W_{ij} = (\epsilon_0\, C_2\, p_i\, p_j)^2 (\epsilon_0\, C_2'\, p_k\, p_m)^2$ $(i, j, k, m = 1, 2, 3, 4;\ i \neq j \neq k \neq m)$; $W_0 = (C_2)^1(\epsilon_0^2\, C_2'\, p_1\, p_2\, p_3\, p_4)^3$; $W_* = (C_2')^1(\epsilon_0^2\, C_2\, p_1\, p_2\, p_3\, p_4)^3$; $W_i = (\epsilon_0\, p_i)^1(\epsilon_0\, C_2\, C_2'\, p_j\, p_k\, p_m)^3$. The factors of each of these surfaces are interchanged by the involution I_{1234}.

The ruled quartic surface $P_K(\epsilon_0^3\, C_2\, C_2')^4$ is the P-surface of the F-curve K, of the first kind, of order 29, 6-fold at p_i $(i = 1, 2, 3, 4)$ and meeting ϵ_0 18 times and C_2, C_2' each 25 times. The surface of coincident points is $R(\epsilon_0^4\, C_2^2\, C_2'^2\, p_1^2\, p_2^2\, p_3^2\, p_4^2)^8$. The involution possesses self-corresponding F-curves of the second kind as follows: the 12 lines ρ on R and P_K and across ϵ_0, C_2 and C_2'; the 4 lines r_i on p_i and across ϵ_0 and C_2; the 4 lines s_i on p_i and across ϵ_0 and C_2'; the line t bisecant to C_2 and C_2'; the 6 rational cubics C_{ij} on p_i, p_j and meeting ϵ_0 twice and C_2, C_2' each 3 times; and the rational quintic C_5 on p_1, p_2, p_3, p_4 and meeting ϵ_0, C_2, C_2' each 4 times.

The homaloidal web H and P-surfaces of the involution have characteristics with respect to the base $(n\colon K, \epsilon_0, C_2, C_2', p_i, p_j, p_k, p_m;\ \rho, r_i, s_i, t, C_{ij}, C_5)$ as follows:

$$\begin{array}{ll} H\ (27\colon 1, 13, 7, 7, 8, 8, 8, 8; 1, 1, 1, 1, 3, 5), & \\ P_K\,(\ 4\colon 0,\ \ 3, 1, 1, 0, 0, 0, 0; 1, 0, 0, 0, 0, 0), & \\ P_{\epsilon_0}(20\colon 1, 10, 5, 5, 6, 6, 6, 6; 1, 1, 1, 0, 2, 4), & \\ P_{C_2}(24\colon 1, 11, 7, 6, 7, 7, 7, 7; 1, 1, 0, 2, 3, 4), & \\ P_{C_2'}(24\colon 1, 11, 6, 7, 7, 7, 7, 7; 1, 0, 1, 2, 3, 4), & \\ P_{p_i}(\ 4\colon 0,\ \ 2, 1, 1, 2, 1, 1, 1; 0, 1, 1, 0, 1, 1) & \qquad (i = 1, 2, 3, 4). \end{array}$$

Let G be the group of Cremona transformations generated by involutions I_{jkmn} with fixed F-curves ϵ_0, C_2, C_2' and variable F-points p_j, p_k, p_m, p_n chosen from the ρ generic points p_i $(i = 1, 2, 3, \cdots, \rho)$, and hence the variable F-curve K.

A surface s_z with characteristic $(z\colon \bar{z}, z_0, z_1, z_2, x_1, x_2, x_3, x_4)$ defined as in the preceding cases is transformed by I_{1234} into a surface $s'_{z'}$. If in the linear transformation giving the characteristic of $s'_{z'}$ we make the substitutions: $u = z - z_0 - z_1 - z_2$, $t_0 = 8z - 6z_0 - 7z_1 - 7z_2$, $t_i = x_i$; then, aside from $u' = \bar{z}$, $u = \bar{z}'$, it becomes the involution

$$i_{1234}\colon\quad \begin{aligned} t_0' &= 5t_0 - 6t_1 - 6t_2 - 6t_3 - 6t_4, \\ t_i' &= t_0 - t_1 - t_2 - t_3 - t_4 - t_i \qquad (i = 1, 2, 3, 4), \\ t_j' &= t_j \qquad (j > 4). \end{aligned}$$

Let $g_\rho(4, 1, 1)$ be the group of linear transformations generated by involutions i_{jkmn} $(j, k, m, n = 1, 2, 3, \cdots, \rho; j \neq k \neq m \neq n)$. It belongs to the type $g_\rho(r, \epsilon, e)$.

The characteristics of the generic element G' of the Cremona group G with respect to the base $(n$: K, ϵ_0, C_2, C_2', p_1, p_2, $\cdots$, $p_\rho)$ and in terms of the coefficients of the corresponding element g' of the group $g_\rho(4, 1, 1)$ are:

$$\begin{array}{llllllll}
H & (32f \pm 5: & 1, & 16f \pm 3, & 8f \pm 1, & 8f \pm 1, & 8\alpha_{10}, 8\alpha_{20}, \cdots, 8\alpha_{\rho 0}), \\
P_{\epsilon_0} & (\ 4\alpha_{00} & : 1, & 2\alpha_{00}, & \alpha_{00}, & \alpha_{00}, & 6\alpha_{10}, 6\alpha_{20}, \cdots, 6\alpha_{\rho 0}), \\
P_{C_2} & (28f \pm 4: & 1, & 14f \pm 3, & 7f, & 7f \pm 1, & 7\alpha_{10}, 7\alpha_{20}, \cdots, 7\alpha_{\rho 0}), \\
P_{C_2'} & (28f \pm 4: & 1, & 14f \pm 3, & 7f \pm 1, & 7f, & 7\alpha_{10}, 7\alpha_{20}, \cdots, 7\alpha_{\rho 0}), \\
P_{p_i} & (\ 4g & : 0, & 2g, & g, & g, & \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i})
\end{array}$$

$(i = 1, 2, \cdots, \rho)$, where $\alpha_{00} = 6f \pm 1$, $\alpha_{0i} = 6g$, f and g integers, and where the variable F-curve of the first kind K corresponds to the fixed P-surface $P_K(\epsilon_0^3 C_2 C_2')^4$. The theorem is proved by showing that $g'i_{1234} \sim G'I_{1234}$, $g'i_{123,\rho+1} \sim G'I_{123,\rho+1}$, $g'i_{12,\rho+1,\rho+2} \sim G'I_{12,\rho+1,\rho+2}$, $g'i_{1,\rho+1,\rho+2,\rho+3} \sim G'I_{1,\rho+1,\rho+2,\rho+3}$ and $g'i_{\rho+1,\rho+2,\rho+3,\rho+4} \sim G'I_{\rho+1,\rho+2,\rho+3,\rho+4}$.

VII. *The involution I_{12} determined by the web of quartic surfaces $W = (\epsilon_0^2\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \epsilon_4\ D^2\ p_1\ p_2)^4$ where the line ϵ_j $(j = 1, 2, 3, 4)$ meets the double line ϵ_0 once and the double point D is in generic position with respect to the rest of the base.* Two members of the web W meet in ϵ_0^4, ϵ_1, ϵ_2, ϵ_3, ϵ_4 and an elliptic octavic curve C_8 on p_1, p_2, 4-fold at D and meeting ϵ_0 4 times and ϵ_j 3 times. The web W contains the 16 uniquely determined degenerate members: $W_{jk}^i = (\epsilon_0\ \epsilon_j\ \epsilon_k\ D\ p_i)^2 (\epsilon_0\ \epsilon_m\ \epsilon_n\ D\ p_h)$ $(j, k, m, n = 1, 2, 3, 4; j \neq k \neq m \neq n; i, h = 1, 2; i \neq h)$; $W_j = (\epsilon_0\ \epsilon_j)^1(\epsilon_0\ \epsilon_k\ \epsilon_m\ \epsilon_n\ D^2\ p_1\ p_2)^3$; $W^j = (\epsilon_j\ D)^1(\epsilon_0^2\ \epsilon_k\ \epsilon_m\ \epsilon_n\ D\ p_1\ p_2)^3$; $W_D = (\epsilon_0\ D)^1(\epsilon_0\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \epsilon_4\ D\ p_1\ p_2)^3$; and $W_{12} = (D\ p_1\ p_2)^1\ (\epsilon_0^2\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \epsilon_4\ D)^3$. The factors of all of these surfaces except W_D and W_{12} are interchanged by the involution.

The plane $P_K(\epsilon_0\ D)^1$ and the ruled cubic surface $P_L(\epsilon_0^2\ \epsilon_1\ \epsilon_2\ \epsilon_3\ \epsilon_4\ D)^3$ are P-surfaces corresponding respectively to the F-curves of the first kind: K, of order 9, 2-fold at p_1, p_2, 3-fold at D, and meeting ϵ_0 twice and ϵ_1, ϵ_2, ϵ_3, ϵ_4 each 5 times; and L, of order 5, 2-fold at p_1, p_2, 3-fold at D and meeting ϵ_0 twice, ϵ_1, ϵ_2, ϵ_3, ϵ_4 each once. The surface of coincident points is $R(\epsilon_0^4\ \epsilon_1^2\ \epsilon_2^2\ \epsilon_3^2\ \epsilon_4^2\ D^4\ p_1^2\ p_2^2)^8$. The involution possesses self-corresponding F-curves of the second kind as follows: the 4 lines ρ on R and P_K across ϵ_0 and D; the 4 conics θ on P_L and R meeting ϵ_0, ϵ_1, ϵ_2, ϵ_3, ϵ_4 each once and on D; the 6 lines r_{jk} on D and across ϵ_j, ϵ_k; the 8 conics C_{ij} on p_i and meeting ϵ_0, ϵ_k, ϵ_m, ϵ_n each once; the rational cubic C_3 on D, p_1, p_2 and meeting ϵ_0 twice and ϵ_1, ϵ_2, ϵ_3, ϵ_4 each once; and the rational quintic C_5 on p_1, p_2, 3-fold at D and meeting ϵ_0, ϵ_1, ϵ_2, ϵ_3, ϵ_4 each twice.

The homaloidal web H and P-surfaces of the involution I_{12} have the following characteristics with respect to the base $(n$: K, L, ϵ_0, ϵ_j, ϵ_k, ϵ_m, ϵ_n, D, p_i, p_h; ρ, r_{jk}, θ, C_{ij}, C_{ik}; C_{im}, C_{in}, C_3, $C_5)$:

$$\begin{array}{l}
H\ (24\colon 1, 2, 11, 6, 6, 6, 6, 13, 8, 8; 1, 1, 2, 2, 2, 2, 2, 3, 5),\\
P_K(\ 1\colon 0, 0,\ \ 1, 0, 0, 0, 0,\ \ 1, 0, 0; 1, 0, 0, 0, 0, 0, 0, 0, 0),\\
P_L(\ 3\colon 0, 0,\ \ 2, 1, 1, 1, 1,\ \ 1, 0, 0; 0, 0, 1, 0, 0, 0, 0, 0, 0),\\
P_{\epsilon_0}(12\colon 1, 1,\ \ 6, 3, 3, 3, 3,\ \ 6, 4, 4; 1, 0, 1, 1, 1, 1, 1, 2, 2),\\
P_{\epsilon_j}(\ 9\colon 0, 1,\ \ 4, 3, 2, 2, 2,\ \ 5, 3, 3; 0, 1, 1, 0, 1, 1, 1, 1, 2),\\
P_D(12\colon 1, 1,\ \ 5, 3, 3, 3, 3,\ \ 7, 4, 4; 1, 1, 1, 1, 1, 1, 1, 1, 3),\\
P_{p_i}(\ 4\colon 0, 0,\ \ 2, 1, 1, 1, 1,\ \ 2, 2, 1; 0, 0, 0, 1, 1, 1, 1, 1, 1).
\end{array}$$

Let G be the group of Cremona transformations generated by involutions I_{ab} with fixed F-curves ϵ_0, ϵ_1, ϵ_2, ϵ_3, ϵ_4 and F-point D and variable F-points p_a, p_b chosen from the ρ generic points p_i $(i = 1, 2, 3, \cdots, \rho)$, and hence the variable F-curves K and L.

Let a surface s_z have the characteristic $(z\colon \bar{z}_1, \bar{z}_2, z_0, z_1, z_2, z_3, z_4, y_0, x_1, x_2)$, where z is the order of s_z and $\bar{z}_1, \bar{z}_2, z_0, z_1, z_2, z_3, z_4, y_0, x_1, x_2$ are its multiplicities on the F-curves K, L, ϵ_0, ϵ_1, ϵ_2, ϵ_3, ϵ_4 and at the F-points D, p_1, p_2, respectively. The characteristic of the transform $s'_{z'}$ of s_z by I_{12} is given by a linear transformation. If in this transformation we make the substitutions: $u_1 = z - z_0 - y_0$, $u_2 = 2z - z_0 - z_1 - z_2 - z_3 - z_4 - y_0$, $t_0 = 8z - 4z_0 - 3z_1 - 3z_2 - 3z_3 - 3z_4 - 4y_0$, $t_i = x_i$; then, aside from $u'_1 = \bar{z}_1$, $u_1 = \bar{z}'_1$, $u'_2 = \bar{z}_2$, $u_2 = \bar{z}'_2$, it becomes the generator i_{12} of Case III. Hence the linear group generated by transformations i_{ab} $(a, b = 1, 2, 3, \cdots, \rho; a \neq b)$ is a $g_\rho(2, 1, 1)$.

The generic element G' of the Cremona group G has the following characteristics with respect to the base $(n\colon K, L, \epsilon_0, \epsilon_j, \epsilon_k, \epsilon_m, \epsilon_n, D, p_1, p_2, \cdots, p_\rho)$ $(j, k, m, n = 1, 2, 3, 4; j \neq k \neq m \neq n)$:

$$\begin{array}{l}
H\ (8\alpha_{00}\colon 1, 2, 4\alpha_{00} \pm 1, 2\alpha_{00}, 2\alpha_{00}, 2\alpha_{00}, 2\alpha_{00}, 4\alpha_{00} \mp 1, 8\alpha_{10}, \cdots, 8\alpha_{\rho 0}),\\
P_{\epsilon_0}(4\alpha_{00}\colon 1, 1, 2\alpha_{00}, \alpha_{00}, \alpha_{00}, \alpha_{00}, \alpha_{00}, 2\alpha_{00}, 4\alpha_{10}, \cdots, 4\alpha_{\rho 0}),\\
P_{\epsilon_j}(3\alpha_{00}\colon 0, 1, 6f \pm 2, 3f, 3f \pm 1, 3f \pm 1, 3f \pm 1, 6f \pm 1, 3\alpha_{10}, \cdots, 3\alpha_{\rho 0}),\\
P_D(4\alpha_{00}\colon 1, 1, 2\alpha_{00} \pm 1, \alpha_{00}, \alpha_{00}, \alpha_{00}, \alpha_{00}, 2\alpha_{00} \mp 1, 4\alpha_{10}, \cdots, 4\alpha_{\rho 0}),\\
P_{p_i}(\alpha_{0i}\colon 0, 0, 2g, g, g, g, g, 2g, \alpha_{1i}, \cdots, \alpha_{\rho i})
\end{array}$$

$(i = 1, 2, 3, \cdots, \rho)$, where $\alpha_{00} = 4f \pm 1$, $\alpha_{0i} = 4g$, f and g integers, and where K and L are variable F-curves of the first kind corresponding respectively to the fixed P-surfaces $P_K(\epsilon_0\, D)^1$ and $P_L(\epsilon_0^2\, \epsilon_1\, \epsilon_2\, \epsilon_3\, \epsilon_4\, D)^3$. The proof is the same as for the theorem of Case III.

VIII. *The involution I_{1234} determined by the web of quartic surfaces $W = (\epsilon_0^2\, \epsilon_1\, \epsilon_2\, D_1^2\, D_2^2\, p_1\, p_2\, p_3\, p_4)^4$ where the double line ϵ_0 meets each of the lines ϵ_1, ϵ_2 once.* Two members of the web W meet in ϵ_0^4, ϵ_1, ϵ_2 and an elliptic 10-ic C_{10} on p_1, p_2, p_3, p_4, 4-fold at D_1, D_2 and meeting ϵ_0 6 times and ϵ_1, ϵ_2 each 3 times. The web contains the 5 uniquely determined degenerate members: $W_j = (\epsilon_0\, \epsilon_j)^1(\epsilon_0\, \epsilon_k\, D_1^2\, D_2^2\, p_1\, p_2\, p_3\, p_4)^3$ $(j, k = 1, 2; j \neq k)$, $W^k = (\epsilon_0\, D_k)^1\, (\epsilon_0\, \epsilon_1\, \epsilon_2\, D_k\, D_j^2\, p_1\, p_2\, p_3\, p_4)^3$,

$W_0 = (\epsilon_0\,\epsilon_1\,\epsilon_2\,D_1\,D_2)^2\,(\epsilon_0\,D_1\,D_2\,p_1\,p_2\,p_3\,p_4)^2$. The factors of W_j are interchanged by the involution.

The quadric $P_L(\epsilon_0\,\epsilon_1\,\epsilon_2\,D_1\,D_2)^2$ is the P-surface of an F-curve L, of the first kind, of order 9, 2-fold at p_1, p_2, p_3, p_4, 3-fold at D_1, D_2 and meeting ϵ_0 6 times and ϵ_1, ϵ_2 each once. The plane $P_{K_k}(\epsilon_0\,D_K)^1$ is the P-surface of an F-curve K_k of the first kind, of order 13, 2-fold at p_1, p_2, p_3, p_4, 3-fold at D_k, 7-fold at D_j and meeting ϵ_0 6 times and ϵ_1, ϵ_2 each 5 times. The surface of coincident points is $R(\epsilon_0^4\,\epsilon_1^2\,\epsilon_2^2\,D_1^4\,D_2^4\,p_1^2\,p_2^2\,p_3^2\,p_4^2)^8$. The involution possesses the following self-corresponding F-curves of the second kind: the 2 sets of 4 lines ρ_k on D_k and across ϵ_0 and on R and P_{K_k}; the line r on D_1 and D_2; the 4 conics θ on D_1, D_2 and meeting ϵ_0, ϵ_1, ϵ_2 each once and on R and P_L; the 8 conics C_{ij} on D_1, D_2 and p_i and meeting ϵ_0, ϵ_j each once; the 6 cubics $C_3^{(ih)}$ $(i, h = 1, 2, 3, 4; i \neq h)$ on D_1, D_2, p_i, p_h and meeting ϵ_0 twice and ϵ_1, ϵ_2 each once; and the rational septimic C_7 on p_1, p_2, p_3, p_4, 3-fold at D_1, D_2 and meeting ϵ_0 4 times and ϵ_1, ϵ_2 each twice.

The homaloidal web H and the P-surfaces of the involution I_{1234} have characteristics with respect to the base $(n\colon L, K_k, K_j, \epsilon_0, \epsilon_j, \epsilon_k, D_k, D_j, p_i, p_h, p_m, p_n;\ \rho_k, r, \theta, C_{ij}, C_3^{(ih)}, C_7)$ $(j, k = 1, 2; j \neq k; i, h, m, n = 1, 2, 3, 4; i \neq h \neq m \neq n)$ as follows:

$$
\begin{array}{llllllllllllllllllll}
H & (33\colon 2, & 1, & 1, & 16, & 8, & 8, & 17, & 17, & 10, & 10, & 10, & 10; & 1, & 1, & 2, & 2, & 3, & 7), \\
P_{K_k} & (\ 1\colon 0, & 0, & 0, & 1, & 0, & 0, & 1, & 0, & 0, & 0, & 0, & 0; & 1, & 0, & 0, & 0, & 0, & 0), \\
P_L & (\ 2\colon 0, & 0, & 0, & 1, & 1, & 1, & 1, & 1, & 0, & 0, & 0, & 0; & 0, & 0, & 1, & 0, & 0, & 0), \\
P_{\epsilon_0} & (20\colon 1, & 1, & 1, & 10, & 5, & 5, & 10, & 10, & 6, & 6, & 6, & 6; & 1, & 0, & 1, & 1, & 2, & 4), \\
P_{\epsilon_j} & (10\colon 1, & 0, & 0, & 5, & 3, & 2, & 5, & 5, & 3, & 3, & 3, & 3; & 0, & 0, & 1, & 1, & 1, & 2), \\
P_{D_k} & (13\colon 1, & 1, & 0, & 6, & 3, & 3, & 7, & 7, & 4, & 4, & 4, & 4; & 1, & 1, & 1, & 1, & 1, & 3), \\
P_{p_i} & (\ 4\colon 0, & 0, & 0, & 2, & 1, & 1, & 2, & 2, & 2, & 1, & 1, & 1; & 0, & 0, & 0, & 1, & 1, & 1).
\end{array}
$$

Let G be the group of Cremona transformations generated by involutions I_{abcd} with fixed F-curves ϵ_0, ϵ_1, ϵ_2 and F-points D_1, D_2 and variable F-points p_a, p_b, p_c, p_d chosen from the ρ generic points p_i $(i = 1, 2, 3, \cdots, \rho)$, and hence the variable F-curves L, K_1, K_2.

If in the linear transformation giving the characteristic of the transform $s_{z'}'$ of a surface s_z with characteristic $(z\colon \bar{z}_0, \bar{z}_1, \bar{z}_2, z_0, z_1, z_2, y_1, y_2, x_1, x_2, x_3, x_4)$ with respect to the involution I_{1234}, we make the substitutions: $u_0 = 2z - z_0 - z_1 - z_2 - y_1 - y_2$, $u_k = z - z_0 - y_k$ $(k = 1, 2)$, $t_0 = 10z - 6z_0 - 3z_1 - 3z_2 - 4y_1 - 4y_2$, $t_i = x_i$; then, aside from $u_0' = \bar{z}_0$, $u_0 = \bar{z}_0'$, $u_k' = \bar{z}_k$, $u_k = \bar{z}_k'$, it becomes the generator i_{1234} of the linear group $g_\rho(4, 1, 1)$ of Case VI. The generic element G' of the Cremona group G has the following characteristics with respect to the base $(n\colon L, K_k, K_j, \epsilon_0, \epsilon_j, \epsilon_k, D_k, D_j, p_1, p_2, p_3, \cdots, p_\rho)$:

$$
\begin{array}{lllllllllll}
H & (40f\pm7\colon & 2,1,1, & 20f\pm4, & 10f\pm2, & 10f\pm2, & 20f\pm3, & 20f\pm3, & 10\alpha_{10}, & \cdots, & 10\alpha_{\rho0}), \\
P_{\epsilon_0} & (\ 4\alpha_{00} & \colon 1,1,1, & 2\alpha_{00}, & \alpha_{00}, & \alpha_{00}, & 2\alpha_{00}, & 2\alpha_{00}, & 6\alpha_{10}, & \cdots, & 6\alpha_{\rho0}), \\
P_{\epsilon_j} & (\ 2\alpha_{00} & \colon 1,0,.0, & \alpha_{00}, & 3f, & 3f\pm1, & \alpha_{00}, & \alpha_{00}, & 3\alpha_{10}, & \cdots, & 3\alpha_{\rho0}),
\end{array}
$$

$$P_{D_k}(16f{\pm}3:\ 1, 1, 0,\ 8f{\pm}2,\ 4f{\pm}1,\ 4f{\pm}1,\ 8f{\pm}1,\ 8f{\pm}1,\ 4\alpha_{10}, \cdots, 4\alpha_{\rho 0}),$$
$$P_{p_i}(\ 4g\ :\ 0, 0, 0,\ 2g\ ,\ g\ ,\ g\ ,\ 2g\ ,\ 2g\ ,\ \alpha_{1i}, \cdots, \alpha_{\rho i}),$$

where $\alpha_{00} = 6f \pm 1$, $\alpha_{0i} = 6g$, f and g integers, and where L, K_1 and K_2 are variable F-curves corresponding to fixed P-surfaces. The proof is the same as for the theorem of Case VI.

IX. *The involution I_{123456} determined by the web of quartic surfaces* $W = (\epsilon_0^2 D_1^2 D_2^2 D_3^2 p_1 p_2 p_3 p_4 p_5 p_6)^4$. Two members of the web W meet in ϵ_0^4 and an elliptic 12-ic C_{12} with 4-fold points at D_1, D_2, D_3, simple points at p_i ($i = 1, 2, 3, 4, 5, 6$) and meeting ϵ_0 8 times. The web W contains the 14 uniquely determined degenerate members: $W_a = (\epsilon_0 D_a)^1(\epsilon_0 D_a D_b^2 D_c^2 p_1 p_2 p_3 p_4 p_5 p_6)^3$ ($a, b, c = 1, 2, 3$; $a \neq b \neq c$), $W_0 = (D_1 D_2 D_3)^1(\epsilon_0^2 D_1 D_2 D_3 p_1 p_2 p_3 p_4 p_5 p_6)^3$, $W_{ihj} = (\epsilon_0 D_1 D_2 D_3 p_i p_h p_j)^2(\epsilon_0 D_1 D_2 D_3 p_k p_m p_n)^2$ ($i, h, j, k, m, n = 1, 2, 3, 4, 5, 6$; $i \neq h \neq j \neq k \neq m \neq n$). The factors of the 10 surfaces W_{ihj} are interchanged by the involution.

The plane $(D_1 D_2 D_3)^1$ is the P-surface of an F-curve L, of the first kind, of order 13, with a double point at p_i, a triple point at D_a and meeting ϵ_0 10 times. The plane $P_{K_a}(\epsilon_0 D_a)^1$ is the P-surface of an F-curve K_a, of the first kind, of order 17, 3-fold at D_a, 7-fold at D_b, D_c, 2-fold at p_i and meeting ϵ_0 10 times. The surface of coincident points is $R(\epsilon_0^4 D_1^4 D_2^4 D_3^4 p_1^2 p_2^2 p_3^2 p_4^2 p_5^2 p_6^2)^8$. The involution possesses self-corresponding F-curves of the second kind as follows: the 3 sets of 4 lines ρ_a on D_a and across ϵ_0 and on P_{K_a} and R; the 4 conics θ on D_1, D_2, D_3 and meeting ϵ_0 once and on P_L and R; the 15 cubics C_{ih} on D_1, D_2, D_3, p_i, and p_h and meeting ϵ_0 twice; and the rational 9-ic C_9 3-fold at D_a, on p_i and meeting ϵ_0 6 times.

The homaloidal web H and the P-surfaces of the involution I_{123456} have characteristics with respect to the base (n: L, K_a, K_b, K_c, ϵ_0, D_a, D_b, D_c, p_i, p_h, p_j, p_k, p_m, p_n ; ρ_a, θ, C_{ih}, C_9) as follows:

$$H\ \ (42:\ 2, 1, 1, 1, 21, 21, 21, 21, 12, 12, 12, 12, 12, 12; 1, 2, 3, 9),$$
$$P_L\ (\ 1:\ 0, 0, 0, 0,\ 0,\ 1,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0; 0, 1, 0, 0),$$
$$P_{K_a}(\ 1:\ 0, 0, 0, 0,\ 1,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0; 1, 0, 0, 0),$$
$$P_{\epsilon_0}(28:\ 1, 1, 1, 1, 14, 14, 14, 14,\ 8,\ 8,\ 8,\ 8,\ 8,\ 8; 1, 1, 2, 6),$$
$$P_{D_a}(14:\ 1, 1, 0, 0,\ 7,\ 7,\ 7,\ 7,\ 4,\ 4,\ 4,\ 4,\ 4,\ 4; 1, 1, 1, 3),$$
$$P_{p_i}(\ 4:\ 0, 0, 0, 0,\ 2,\ 2,\ 2,\ 2,\ 2,\ 1,\ 1,\ 1,\ 1,\ 1; 0, 0, 1, 1).$$

Let G be the group of Cremona transformations generated by involutions I_{ghjkmn} with fixed F-curve ϵ_0 and F-points D_1, D_2, D_3 and variable F-points p_g, p_h, p_j, p_k, p_m, p_n chosen from the ρ generic points p_i ($i = 1, 2, 3, \cdots, \rho$) and hence the variable F-curves L, K_1, K_2, K_3.

If in the linear transformation giving the characteristic of the transform $s'_{z'}$ of the surface s_z, whose characteristic with respect to the involution I_{123456} is (z: $\bar{z}_0$, $\bar{z}_1$, $\bar{z}_2$, $\bar{z}_3$, z_0, y_1, y_2, y_3, x_1, x_2, x_3, x_4, x_5, x_6), we make the substitu-

tions: $u_0 = 2z - z_0 - y_1 - y_2 - y_3$, $u_a = z - z_0 - y_a$ $(a = 1, 2, 3)$, $t_0 = 12z - 8z_0 - 4y_1 - 4y_2 - 4y_3$, $t_i = x_i$; then, aside from $u_0' = \bar{z}_0$, $u_0 = \bar{z}_0'$, $u_a' = \bar{z}_a$, $u_a = \bar{z}_a'$, it becomes:

$$\begin{aligned} & t_0' = 7t_0 - 8t_1 - 8t_2 - 8t_3 - 8t_4 - 8t_5 - 8t_6, \\ i_{123456}: \quad & t_i' = t_0 - t_1 - t_2 - t_3 - t_4 - t_5 - t_6 - t_i \qquad (i = 1, \cdots, 6), \\ & t_h' = t_h \qquad (h > 6). \end{aligned}$$

Let $g_\rho(6, 1, 1)$ be the group of linear transformations generated by involutions i_{ghjkmn} $(g, h, j, k, m, n = 1, 2, 3, \cdots, \rho; g \neq h \neq j \neq k \neq m \neq n)$. $g_\rho(6, 1, 1)$ is evidently of the type $g_\rho(r, \epsilon, e)$. The generic element G' of the Cremona group G has the following characteristics, in terms of the coefficients of the corresponding element g' of the group $g_\rho(6, 1, 1)$, and with respect to the base $(n: L, K_a, K_b, K_c, \epsilon_0, D_a, D_b, D_c, p_1, p_2, p_3, \cdots, p_\rho)$:

$$\begin{array}{ll} H & (6\alpha_{00} : 2, 1, 1, 1, 3\alpha_{00}, 3\alpha_{00}, 3\alpha_{00}, 3\alpha_{00}, 12\alpha_{10}, 12\alpha_{20}, \cdots, 12\alpha_{\rho 0}), \\ P_{\epsilon_0} & (4\alpha_{00} : 1, 1, 1, 1, 2\alpha_{00}, 2\alpha_{00}, 2\alpha_{00}, 2\alpha_{00}, 8\alpha_{10}, 8\alpha_{20}, \cdots, 8\alpha_{\rho 0}), \\ P_{D_a} & (2\alpha_{00} : 1, 1, 0, 0, \alpha_{00}, \alpha_{00}, \alpha_{00}, \alpha_{00}, 4\alpha_{10}, 4\alpha_{20}, \cdots, 4\alpha_{\rho 0}), \\ P_{p_i} & (2f : 0, 0, 0, 0, f, f, f, f, \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i}) \end{array}$$

$(i = 1, 2, 3, \cdots, \rho)$, where $\alpha_{0i} = 4f$, f an integer, and where L, K_1, K_2, K_3 are variable F-curves of the first kind corresponding to fixed P-surfaces. The theorem is proved in the same way as the theorems in the other cases.

X. *The involution I_{12} determined by the web of quartic surfaces $W = (C_2^2 C_4 p_1 p_2)^4$ where the elliptic quartic curve C_4 meets the double conic C_2 4 times.* Two members of the web W meet in C_2^4, C_4 and an elliptic quartic C_4' on p_1, p_2 and meeting C_2, C_4 each 4 times. The web W contains 2 nets of degenerate members: $W_1 = (C_2 C_4)^2(C_2 p_1 p_2)^2$, $W_2 = (C_2)^1(C_2 C_4 p_1 p_2)^3$. W_1 and W_2 have in common the pencil of degenerate members $W_{12} = (C_2)^1(C_2 C_4)^2(p_1 p_2)^1$.

In each plane of the pencil $(p_1 p_2)^1$ the web W determines an involution which is the transform under a quadratic transformation T_2 of a special Geiser's involution having 6 of its 7 base points on a conic. But the curves of the homaloidal net and the p-curves of this involution are generic sections of the surfaces of the homaloidal web H and the P-surfaces of the involution determined in space by the web W. The plane $P_{P_0}(C_2)^1$ is the P-surface of an isolated F-point P_0 on the line joining p_1 and p_2. The involution possesses self-corresponding F-curves of the second kind as follows: the 2 sets of 4 lines r_i $(i = 1, 2)$ on p_i and across C_2 and C_4; and the 2 conics C on p_1, p_2 and meeting C_2, C_4 each twice.

The homaloidal web H and the P-surfaces of the involution I_{12} have characteristics with respect to the base $(n: P_0, C_2, C_4, p_i, p_h; r_i, C)$ $(i, h = 1, 2; i \neq h)$ as follows:

$$H(10\colon 1, 5, 2, 4, 4; 1, 2),$$
$$P_{P_0}(1\colon 0, 1, 0, 0, 0; 0, 0), \qquad P_{C_2}(10\colon 2, 5, 2, 4, 4; 1, 2),$$
$$P_{C_4}(8\colon 0, 4, 1, 4, 4; 1, 2), \qquad P_{p_i}(\ 4\colon 0, 2, 1, 2, 1; 1, 1).$$

Let G be the group of Cremona transformations generated by involutions I_{jk} with fixed F-curves C_2, C_4 and variable F-points p_j, p_k chosen from the ρ generic points p_i $(i = 1, 2, \cdots, \rho)$, and hence the variable F-point P_0.

If in the linear transformation giving the characteristic of the transform $s'_{z'}$ of a surface s_z, having the characteristic $(z\colon \bar{z}, z_0, z_1, x_1, x_2)$ with respect to the involution I_{12}, we make the substitutions: $u = z - 2z_0$, $t_0 = 4z - 4z_0 - 4z_1$, $t_i = x_i$; then, apart from $u' = \bar{z}$, $u = \bar{z}'$, it becomes the generator i_{12} of Case III. Hence the group generated by involutions i_{jk} $(j, k = 1, 2, 3, \cdots, \rho; j \neq k)$ is $g_\rho(2, 1, 1)$. The generic element G' of the Cremona group G has the following characteristics with respect to the base $(n\colon P_0, C_2, C_4, p_1, p_2, p_3, \cdots, p_\rho)$:

$$\begin{array}{ll} H & (4\alpha_{00} - 2\colon 1, 2\alpha_{00} - 1, \alpha_{00} - 1, 4\alpha_{10}, 4\alpha_{20}, \cdots, 4\alpha_{\rho 0}), \\ P_{C_2} & (4\alpha_{00} - 2\colon 2, 2\alpha_{00} - 1, \alpha_{00} - 1, 4\alpha_{10}, 4\alpha_{20}, \cdots, 4\alpha_{\rho 0}), \\ P_{C_4} & (4\alpha_{00} - 4\colon 0, 2\alpha_{00} - 2, \alpha_{00} - 2, 4\alpha_{10}, 4\alpha_{20}, \cdots, 4\alpha_{\rho 0}), \\ P_{p_i} & (\ \alpha_{0i} \qquad \colon 0, 2f \qquad , f \qquad , \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i}) \end{array}$$

$(i = 1, 2, 3, \cdots, \rho)$, where $\alpha_{0i} = 4f$, f an integer, and where P_0 is a variable F-point corresponding to the fixed P-surface $P_{P_0}(C_2)^1$. The theorem is proved as in Case III.

XI. *The involution I_{123} determined by the web of quartic surfaces $W = (C_2^2\, C_3\, p_1\, p_2\, p_3)^4$ where the rational cubic C_3 meets the double conic C_2 3 times.* Two members of the web W meet in C_2^4, C_3 and an elliptic quintic C_5 on p_1, p_2, p_3 and meeting C_2, C_3 each 5 times. W contains the net of degenerate members $W_1 = (C_2)^1 (C_2\, C_3\, p_1\, p_2\, p_3)^3$ and the pencil of degenerate members $W_2 = (C_2\, C_3)^2(C_2\, p_1\, p_2\, p_3)^2$. W_1 and W_2 have in common the uniquely determined member $W_{12} = (C_2)^1 (C_2\, C_3)^2(p_1\, p_2\, p_3)^1$.

The plane $P_{P_0}(C_2)^1$ is the P-surface of an isolated F-point P_0. The quadric $P_K(C_2\, C_3)^2$ is the P-surface of an F-curve K, of the first kind, of order 2, on p_1, p_2, p_3 and meeting C_2 twice. The surface of coincident points is $R(C_2^4\, C_3^2\, p_1^2\, p_2^2\, p_3^2)^8$. The involution possesses self-corresponding F-curves of the second kind as follows: the 2 lines ρ on R and P_K, across C_2 and bisecant to C_3; the 3 sets of 3 lines r_i on p_i and across C_2 and C_3; the 3 conics C_i on p_j, p_k and meeting C_2, C_3 each twice; and the rational cubic C_3' on p_1, p_2, p_3 and meeting C_2, C_3 each 3 times.

The homaloidal web H and the P-surfaces of the involution have the following characteristics with respect to the base $(n\colon P_0, K, C_2, C_3, p_i, p_j, p_k; \rho, r_i, C_i, C_j, C_k, C_3')$:

$$\begin{aligned}
&H\ (16\colon 1, 1, 8, 4, 5, 5, 5; 1, 1, 2, 2, 2, 3),\\
&P_{P_0}(\ 1\colon 0, 0, 1, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0),\\
&P_K(\ 2\colon 0, 0, 1, 1, 0, 0, 0; 1, 0, 0, 0, 0, 0),\\
&P_{C_2}(16\colon 2, 1, 8, 4, 5, 5, 5; 1, 1, 2, 2, 2, 3),\\
&P_{C_3}(16\colon 0, 2, 8, 4, 5, 5, 5; 2, 1, 2, 2, 2, 3),\\
&P_{p_i}(\ 4\colon 0, 0, 2, 1, 2, 1, 1; 0, 1, 0, 1, 1, 1),
\end{aligned}$$

$(i, j, k = 1, 2, 3; i \neq j \neq k)$. Let G be the group of Cremona transformations generated by involutions I_{hjk} with fixed F-curves C_2, C_3 and variable F-points p_h, p_j, p_k chosen from the ρ generic points p_i $(i = 1, 2, 3, \cdots, \rho)$, and hence the variable F-point P_0 and F-curve K.

If in the linear transformation giving the characteristic of the transform $s'_{z'}$ of a surface s_z, having the characteristic $(z\colon \bar{z}_0, \bar{z}_1, z_0, z_1, x_1, x_2, x_3)$ with respect to the involution I_{123}, we make the substitutions: $u_0 = z - 2z_0$, $u_1 = z - z_0 - 2z_1$, $t_0 = 5z - 5z_0 - 5z_1$, $t_i = x_i$; then, apart from $u'_m = \bar{z}_m$, $u_m = \bar{z}'_m$ $(m = 0, 1)$, it becomes the generator i_{123} of Case V. Hence the group generated by involutions i_{hjk} $(h, j, k = 1, 2, 3, \cdots, \rho;\ h \neq j \neq k)$ is $g_\rho(3, 1, 1)$. The generic member G' of the Cremona group G has characteristics with respect to the base $(n\colon P_0, K, C_2, C_3, p_1, p_2, \cdots, p_\rho)$ as follows:

$$\begin{aligned}
&H\ (4\alpha_{00} : 1, 1, 2\alpha_{00}, \alpha_{00}, 5\alpha_{10}, 5\alpha_{20}, \cdots, 5\alpha_{\rho 0}),\\
&P_{C_2}(4\alpha_{00} : 2, 1, 2\alpha_{00}, \alpha_{00}, 5\alpha_{10}, 5\alpha_{20}, \cdots, 5\alpha_{\rho 0}),\\
&P_{C_3}(4\alpha_{00} : 0, 2, 2\alpha_{00}, \alpha_{00}, 5\alpha_{10}, 5\alpha_{20}, \cdots, 5\alpha_{\rho 0}),\\
&P_{p_i}(\ 4f\ : 0, 0, 2f\ , f\ , \alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{\rho i})
\end{aligned}$$

$(i = 1, 2, 3, \cdots, \rho)$, where $\alpha_{0i} = 5f$, f an integer, and where P_0 and K are variable F-point and F-curve corresponding respectively to the fixed P-surfaces $P_{P_0}(C_2)^1$ and $P_K(C_2\,C_3)^2$. The proof is the same as for the theorem of Case V.

XII. *The involution I_{1234} determined by the web of quartic surfaces $W = (C_2^2\ \epsilon_1\ \epsilon_2\ p_1\ p_2\ p_3\ p_4)^4$ where each of the lines ϵ_1, ϵ_2 meets the double conic C_2 once.* Two members of the web W meet in C_2^4, ϵ_1, ϵ_2 and an elliptic sextic C_6 on p_1, p_2, p_3, p_4 and meeting C_2 6 times and ϵ_1, ϵ_2 each 3 times. W contains the net of degenerate members $W_1 = (C_2)^1(C_2\ \epsilon_1\ \epsilon_2\ p_1\ p_2\ p_3\ p_4)^3$ and the 7 uniquely determined degenerate members: $W_0 = (C_2\ \epsilon_1\ \epsilon_2)^2(C_2\,p_1\,p_2\,p_3\,p_4)^2$, $W_{hi} = (C_2\ \epsilon_1\ p_h\ p_i)^2 (C_2\ \epsilon_2\ p_j\ p_k)^2$ $(h, i, j, k = 1, 2, 3, 4; h \neq i \neq j \neq k)$. The factors of the surfaces W_{hi} are interchanged by the involution.

The plane $P_{P_0}(C_2)^1$ is the P-surface of an isolated F-point P_0. The quadric $P_K(C_2\ \epsilon_1\ \epsilon_2)^2$ is the P-surface of an F-curve K, of the first kind, of order 6, on p_1, p_2, p_3, p_4 each twice and meeting C_2 6 times and ϵ_1, ϵ_2 each once. The surface of coincident points is $R(C_2^4\ \epsilon_1^2\ \epsilon_2^2\ p_1^2\ p_2^2\ p_3^2\ p_4^2)^8$. The involution possesses self-corresponding F-curves of the second kind as follows: the 4 lines ρ on R and P_K and across C_2, ϵ_1 and ϵ_2; the 4 sets of 2 lines each r_{im} on p_i and across C_2,

ϵ_m; the 6 conics C_{ij} on p_i, p_j and meeting C_2 twice and ϵ_1, ϵ_2 each once; and the rational quartic C_4 on p_1, p_2, p_3, p_4 and meeting C_2 4 times and ϵ_1, ϵ_2 each twice.

The homaloidal web H and the P-surfaces of the involution I_{1234} have characteristics with respect to the base (n: P_0, K, C_2, ϵ_m, ϵ_n, p_i, p_h, p_j, p_k; ρ, r_{im}, C_{ij}, C_4) as follows:

$$\begin{array}{ll}
H & (20:\ 1, 1, 10, 5, 5, 6, 6, 6, 6; 1, 1, 2, 4),\\
P_{P_0} & (\ 1:\ 0, 0,\ \ 1, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0),\\
P_K & (\ 2:\ 0, 0,\ \ 1, 1, 1, 0, 0, 0, 0; 1, 0, 0, 0),\\
P_{C_2} & (20:\ 2, 1, 10, 5, 5, 6, 6, 6, 6; 1, 1, 2, 4),\\
P_{\epsilon_m} & (10:\ 0, 1,\ \ 5, 3, 2, 3, 3, 3, 3; 1, 1, 1, 2),\\
P_{p_i} & (\ 4:\ 0, 0,\ \ 2, 1, 1, 2, 1, 1, 1; 0, 1, 1, 1),
\end{array}$$

($i, h, j, k = 1, 2, 3, 4$; $i \neq j \neq h \neq k$; $m, n = 1, 2$; $m \neq n$). Let G be the group of Cremona transformations generated by involutions I_{abcd} with fixed F-curves C_2, ϵ_1, ϵ_2, variable F-points p_a, p_b, p_c, p_d selected from the ρ generic points p_i ($i = 1, 2, 3, \cdots, \rho$), and hence the variable F-point P_0 and F-curve K.

If in the linear transformation giving the characteristic of the transform $s'_{z'}$ of a surface s_z, having the characteristic (z: $\bar{z}_0$, $\bar{z}_1$, z_0, z_1, z_2, x_1, x_2, x_3, x_4) with respect to the involution I_{1234}, we make the substitutions: $u_0 = z - 2z_0$, $u_1 = z - z_0 - z_1 - z_2$, $v = z_1 - z_2$, $t_0 = 6z - 6z_0 - 3z_1 - 3z_2$, $t_i = x_i$; then, apart from $u'_s = \bar{z}_s$, $u_s = \bar{z}'_s$ ($s = 0, 1$), $v' = -v$, it becomes the generator i_{1234} of the group $g_\rho(4, 1, 1)$ of Case VI. The characteristics of the generic element G' of the Cremona group G with respect to the base (n: P_0, K, C_2, ϵ_m, ϵ_n, p_1, $\cdots$, p_ρ) are as follows:

$$\begin{array}{ll}
H & (4\alpha_{00}:\ 1, 1, 2\alpha_{00},\ \alpha_{00},\ \alpha_{00}\ \ , 6\alpha_{10}, 6\alpha_{20}, \cdots, 6\alpha_{\rho 0}),\\
P_{C_2} & (4\alpha_{00}:\ 2, 1, 2\alpha_{00},\ \alpha_{00},\ \alpha_{00}\ \ , 6\alpha_{10}, 6\alpha_{20}, \cdots, 6\alpha_{\rho 0}),\\
P_{\epsilon_m} & (2\alpha_{00}:\ 0, 1,\ \ \alpha_{00}, 3f\ \ , 3f\pm 1, 3\alpha_{10}, 3\alpha_{20}, \cdots, 3\alpha_{\rho 0}),\\
P_{p_i} & (4g\ \ :\ 0, 0, 2g\ \ ,\ g\ \ ,\ g\ \ \ ,\ \alpha_{1i},\ \alpha_{2i}, \cdots,\ \alpha_{\rho i}),
\end{array}$$

($i = 1, 2, 3, \cdots, \rho$; $m, n = 1, 2$; $m \neq n$), where $\alpha_{00} = 6f \pm 1$, $\alpha_{0i} = 6g$, f and g integers, and where P_0 and K are variable F-point and F-curve corresponding to fixed P-surfaces.

XIII. *The involution* I_{123456} *determined by the web of quartic surfaces* $W = (C_2^2\, D^2\, p_1\, p_2\, p_3\, p_4\, p_5\, p_6)^4$. Two members of the web W meet in C_2^4 and an elliptic octavic C_8 on p_i ($i = 1, 2, 3, 4, 5, 6$), 4-fold at D and meeting C_2 8 times. W contains the net of degenerate members $W_0 = (C_2)^1(C_2\, D^2\, p_1\, p_2\, p_3\, p_4\, p_5\, p_6)^3$ and the 10 uniquely determined degenerate members $W_{ihj} = (C_2\, D\, p_i\, p_h\, p_j)^2(C_2\, D\, p_k\, p_m\, p_n)^2$ ($i, h, j, k, m, n = 1, 2, 3, 4, 5, 6$; $i \neq h \neq j \neq k \neq m \neq n$). The factors of the surfaces W_{ihj} are interchanged by the involution.

The plane $P_{P_0}(C_2)^1$ is the P-surface of an isolated F-point P_0. The quadric

cone $P_K(C_2 D^2)^2$ is the P-surface of an F-curve K, of the first kind, of order 20, 4-fold at p_i ($i = 1, 2, 3, 4, 5, 6$), 6-fold at D and meeting C_2 20 times. The surface of coincident points is $R(C_2^4 D^4 p_1^2 p_2^2 p_3^2 p_4^2 p_5^2 p_6^2)^3$. The involution possesses self-corresponding F-curves of the second kind as follows: the 8 lines ρ on D and across C_2 and on P_K and R; the 15 conics C_{ij} on D, p_i, p_j and meeting C_2 twice; and the rational sextic C_6 on $p_1, p_2, p_3, p_4, p_5, p_6$, 3-fold at D and meeting C_2 6 times.

The homaloidal web H and P-surfaces of the involution I_{123456} have characteristics with respect to the base (n: $P_0, K, C_2, D, p_i, p_h, p_j, p_k, p_m, p_n$; ρ, C_{ij}, C_6) as follows:

$$
\begin{aligned}
&H \ (28\colon 1, 1, 14, 14, 8, 8, 8, 8, 8, 8; 1, 2, 6),\\
&P_{P_0}(\ 1\colon 0, 0,\ 1,\ 0, 0, 0, 0, 0, 0, 0; 0, 0, 0),\\
&P_K\ (\ 2\colon 0, 0,\ 1,\ 2, 0, 0, 0, 0, 0, 0; 1, 0, 0),\\
&P_{C_2}(28\colon 2, 1, 14, 14, 8, 8, 8, 8, 8, 8; 1, 2, 6),\\
&P_D\ (14\colon 0, 1,\ 7,\ 7, 4, 4, 4, 4, 4, 4; 1, 1, 3),\\
&P_{p_i}(\ 4\colon 0, 0,\ 2,\ 2, 2, 1, 1, 1, 1, 1; 0, 1, 1).
\end{aligned}
$$

Let G be the group of Cremona transformations generated by involutions I_{ghjkmn} with fixed F-curve C_2 and F-point D and variable F-points $p_g, p_h, p_j, p_k, p_m, p_n$ chosen from the ρ generic points p_i ($i = 1, 2, 3, \cdots, \rho$) and hence the variable F-point P_0 and F-curve K.

If in the linear transformation giving the characteristic of the transform $s'_{z'}$ of a surface s_z, with characteristic (z: $\bar{z}_0, \bar{z}_1, z_0, y_0, x_1, x_2, x_3, x_4, x_5, x_6$) with respect to the involution I_{123456}, we make the substitutions: $u_0 = z - 2z_0$, $u_1 = z - z_0 - y_0$, $t_0 = 8z - 8z_0 - 4y_0$, $t_i = x_i$; then, apart from $u'_0 = \bar{z}_0$, $u_0 = \bar{z}'_0$, $u'_1 = \bar{z}_1$, $u_1 = \bar{z}'_1$, it becomes the generator i_{123456} of the linear group $g_\rho(6, 1, 1)$ of Case IX. The characteristics of the generic element G' of the Cremona group G with respect to the base (n: $P_0, K, C_2, D, p_1, p_2, p_3, \cdots, p_\rho$) are as follows:

$$
\begin{aligned}
&H\ (4\alpha_{00} : 1, 1, 2\alpha_{00}, 2\alpha_{00}, 8\alpha_{10}, 8\alpha_{20}, 8\alpha_{30}, \cdots, 8\alpha_{\rho 0}),\\
&P_{C_2}(4\alpha_{00} : 2, 1, 2\alpha_{00}, 2\alpha_{00}, 8\alpha_{10}, 8\alpha_{20}, 8\alpha_{30}, \cdots, 8\alpha_{\rho 0}),\\
&P_D\ (2\alpha_{00} : 0, 1, \alpha_{00}, \alpha_{00}, 4\alpha_{10}, 4\alpha_{20}, 4\alpha_{30}, \cdots, 4\alpha_{\rho 0}),\\
&P_{p_i}(2f\ : 0, 0, f, f, \alpha_{1i}, \alpha_{2i}, \alpha_{3i}, \cdots, \alpha_{\rho i})
\end{aligned}
$$

($i = 1, 2, 3, \cdots, \rho$), where $\alpha_{0i} = 4f$, f an integer, and where P_0 and K are variable F-point and F-curve corresponding to fixed P-surfaces.

Conclusion. We have thus exhibited a rather large aggregate of groups of Cremona transformations simply isomorphic to the linear groups $g_\rho(\alpha)$ and $g_\rho(r, \epsilon, e)$.

UNIVERSITY OF ILLINOIS.

WEAKLY COMPLETE BANACH SPACES

By H. H. Goldstine

Banach has shown that if every bounded sequence of elements from a separable Banach space contains a subsequence which converges weakly to some element of the space, the space is equivalent to the adjoint of its adjoint and recently the converse of this theorem has been shown to be true.[1] In this paper we shall investigate necessary and sufficient conditions that a space be equivalent to the adjoint of its adjoint. We show that if the adjoint space is separable, such a condition is that the space be weakly complete. In fact, the assumption of separability can be dispensed with by redefining the notion of weak completeness.

Most of the results herein contained follow from a representation theorem that is based upon one of Hildebrandt's.[2] This result enables us to examine the analytic character of functionals in the adjoint of the adjoint of a Banach space.

1. **Definitions. The representation theorems.** Before proceeding to establish the basic theorems of this section, we shall find it convenient to fix upon certain notations and definitions which will be used throughout this paper. For this purpose we agree that

(a) $\mathfrak{R}$ is the set of all real numbers;

(b) $\mathfrak{P}$ is an arbitrary class of elements p;

(c) $\mathfrak{Q}$ is a non-null subset of $\mathfrak{P}$;

(d) $\mathfrak{X}$ is a linear class of functionals ξ on $\mathfrak{P}$ to $\mathfrak{R}$, which are bounded on $\mathfrak{Q}$; i.e., to each ξ there corresponds a number n_ξ such that

$$|\xi(q)| \leqq n_\xi \qquad (q \in \mathfrak{Q});$$

(e) the norm, $\|\xi\|$, of a functional ξ in $\mathfrak{X}$ is the least upper bound of $|\xi(q)|$ for q in $\mathfrak{Q}$;

(f) $\mathfrak{D}$ is the class of *all* partitions or divisions δ of $\mathfrak{Q}$ into a finite number of mutually exclusive subsets E_i. In accordance with Hildebrandt's notation we say that $\delta_1 \geqq \delta_2$ in the event that every set E_1 of δ_1 is contained in some subset E_2 of δ_2;

Received October 25, 1937; presented to the American Mathematical Society, November 27, 1937.

[1] See S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 189, and V. Gantmakher and V. Šmulian, *Sur les espaces linéaires dont la sphère unitaire est faiblement compacte*, Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS, vol. 17 (1937), pp. 91–94.

[2] *On bounded linear functional operations*, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 868–875.

(g) if a_δ is a real-valued function defined on $\mathfrak{D}$, then $\lim_\delta a_\delta = a$ in case to each positive number e there corresponds a δ_e in $\mathfrak{D}$ such that $| a_\delta - a | < e$, whenever $\delta \geqq \delta_e$. This limit is utilized by Hildebrandt and is a special case of the general limit of E. H. Moore and H. L. Smith.[3]

We now proceed to establish two representation theorems which are essentially due to Hildebrandt.[4]

THEOREM I. *Every linear and continuous functional M on $\mathfrak{X}$ to $\mathfrak{R}$ is expressible as*

$$M(\xi) = \int \xi \, d\lambda \qquad (\xi \in \mathfrak{X}),$$

where λ is some real-valued, additive function of bounded variation[5] defined on the class of all subsets E of $\mathfrak{Q}$. The integral is either of the Hildebrandt-Stieltjes or of the Hildebrandt-Lebesgue type. Moreover, the bound or norm of M, $\| M \|$, is the total variation of λ.

In order to prove this theorem let $\mathfrak{Y}$ be the class of *all* functionals η on $\mathfrak{P}$ to $\mathfrak{R}$, that are bounded on $\mathfrak{Q}$, and let the norm $\| \eta \|$ of η be the upper bound of $| \eta |$ on $\mathfrak{Q}$. Clearly $\mathfrak{X}$ is a linear subset of $\mathfrak{Y}$. Two functionals η_1, η_2 of $\mathfrak{Y}$ are said to be equivalent when they are equal on $\mathfrak{Q}$; i.e., for every q, $\eta_1(q) = \eta_2(q)$. Then if η_1 is equivalent to η_2, the norm of η_1 equals the norm of η_2 and the norm of $\eta_1 - \eta_2$ is zero. Moreover, it is also clear that this relation is reflexive, symmetric and transitive. Consequently it divides $\mathfrak{Y}$ into mutually exclusive and exhaustive subsets $[\eta]$; hence $[\eta]$ consists of all η_1 that are equivalent to η. The class of all these subsets becomes a normed linear space when the following definitions are made: the norm of $[\eta]$ is the norm of η; $a[\eta]$ is the set $[a\eta]$; and $[\eta_1] + [\eta_2] = [\eta_1 + \eta_2]$. It is evident that these operations are well defined.

If $\delta = (E_1, \cdots, E_n)$ is any partition in $\mathfrak{D}$, q_i is any point in E_i $(i = 1, 2, \cdots, n)$, and if η is any functional in $\mathfrak{Y}$, then η_δ is defined to be the sum

$$\sum_{i=1}^{n} \eta(q_i)\varphi_{E_i},$$

where φ_E is the characteristic function of the subset E of $\mathfrak{Q}$; i.e., $\varphi_E(p)$ is unity if p is in E and is zero if p is not. It is, of course, true that η_δ and φ_E are in $\mathfrak{Y}$.

The limit as to δ of η_δ exists and is equivalent to η. Moreover, the limit, $\lim_\delta [\eta_\delta]$, exists and is $[\eta]$, both limits being taken in the sense of the norm. The proof of the first assertion is entirely analogous to the one given by Hildebrandt,[6] and the second follows at once from this and the definition of the equivalence relation.

[3] *A general form of limit*, American Journal of Mathematics, vol. 44 (1922), pp. 102–121.

[4] *Bounded functional operations*, p. 875.

[5] Ibid., p. 869.

[6] *Bounded functional operations*, p. 870.

Let $N([\xi])$ be defined as $M(\xi)$. Then N is a single-valued function on the class of all subsets $[\xi]$, where ξ is in $\mathfrak{X}$; for, if ξ_1 and ξ_2 are both in $[\xi]$, then

$$| M(\xi_1) - M(\xi_2) | \leqq \| M \| \cdot \| \xi_1 - \xi_2 \| = 0.$$

Furthermore, it is linear, continuous and has its bound $\| N \|$ equal to $\| M \|$. Hence by Banach's extension theorem[7] there exists a linear continuous functional P on the class of all $[\eta]$, having the same bound as N and coinciding with N on the range of definition of N.

Having thus extended the range of N, we define $\lambda(E)$ to be $P([\varphi_E])$ and proceed in the manner of Hildebrandt.[8] That is to say, we can see that

$$P([\xi_\delta]) = \sum_i \xi(q_i)\lambda(E_i);$$

and consequently by means of the continuity of P that

$$M(\xi) = N([\xi]) = \lim_\delta \sum_i \xi(q_i)\lambda(E_i) = \int \xi \, d\lambda.$$

This integral is of the Hildebrandt-Stieltjes type, but since ξ is measurable relative to the class of all subsets E of $\mathfrak{Q}$, in the sense that the set $E\,[a < \xi(q) \leqq b]$ is a subset of $\mathfrak{Q}$, the Hildebrandt-Lebesgue integral exists and equals the Stieltjes integral.[9]

THEOREM II. *Every linear and continuous functional M on $\mathfrak{X}$ to $\mathfrak{R}$ is expressible in the form*

$$M(\xi) = \lim_\delta \sum_p \xi(p)\beta_\delta(p) \qquad (\xi \in \mathfrak{X}),$$

the function β_δ being different from zero for at most a finite number of places p, and

$$\| M \| = \lim_\delta \sum_p | \beta_\delta(p) |.$$

To establish this result it suffices to modify the proof of Hildebrandt[10] as was done in the preceding theorem.

2. **Adjoint spaces.** Throughout the remaining sections it will be understood that

(h) $\mathfrak{P}$ is a normed linear space;
(i) $\mathfrak{Q}$ is the unit sphere [all p such that $\| p \| \leqq 1$] in $\mathfrak{P}$;
(j) $\mathfrak{X}$ is the set of *all* linear continuous functionals ξ on $\mathfrak{P}$ to $\mathfrak{R}$;
(k) $\mathfrak{M}$ is the set of *all* linear continuous functionals M on $\mathfrak{X}$ to $\mathfrak{R}$.

[7] *Opérations Linéaires*, p. 55.
[8] *Bounded functional operations*, pp. 870–872.
[9] Ibid., p. 869.
[10] Ibid., p. 871.

THEOREM III. *To every linear continuous functional M on $\mathfrak{X}$ to $\mathfrak{R}$ there corresponds a function p_δ, defined on $\mathfrak{D}$ and having functional values in $\mathfrak{P}$ such that*

$$M(\xi) = \lim_\delta \xi(p_\delta) \qquad (\xi \epsilon \mathfrak{X}),$$

and

$$\overline{\lim_\delta} \| p_\delta \| < +\infty.$$

Since each ξ is linear,

$$\sum_p \xi(p)\beta_\delta(p) = \xi(\sum_p p\beta_\delta(p));$$

whence if p_δ is defined to be $\sum_p p\beta_\delta(p)$, then

$$\xi(\sum_p p\beta_\delta(p)) = \xi(p_\delta).$$

Moreover, it is evident from the definition of $\mathfrak{Q}$ and from the fact that $\beta_\delta(p) = 0$ if p is not in $\mathfrak{Q}$ that

$$\| p_\delta \| \leqq \sum_q |\beta_\delta(q)| \leqq \| M \|$$

for every δ.

Let us now turn attention to an arbitrary class $\mathfrak{L}$ and a relation R on $\mathfrak{L}\mathfrak{L}$ which is transitive and compositive.[11] We shall say that $\mathfrak{P}$ is *weakly complete relative to* $\mathfrak{L}$ in the event that to every function p_l on $\mathfrak{L}$ to $\mathfrak{P}$ such that

$$\overline{\lim_l} \| p_l \| < +\infty \quad \text{and} \quad \lim_l \xi(p_l) \qquad (\xi \epsilon \mathfrak{X})$$

exists there corresponds a p_0 in $\mathfrak{P}$ with the property that

$$\xi(p_0) = \lim_l \xi(p_l) \qquad (\xi \epsilon \mathfrak{X}).$$

If $\mathfrak{L}$ is the set of positive integers and R is the "greater than" relation, then the restriction that the upper limit of $\| p_l \|$ is finite is a consequence of the existence of the $\lim_l \xi(p_l)$, for every ξ. However, this is not necessarily the case for more general $\mathfrak{L}$R-systems. If $\mathfrak{P}$ is weakly complete relative to every $\mathfrak{L}$ and R, then it is termed *weakly complete.* For separable spaces this definition is equivalent to the one of Banach,[12] as will be shown later. First we shall prove

THEOREM IV. *Every functional M in $\mathfrak{M}$ is uniquely expressible as*

$$M(\xi) = \xi(p_0) \qquad (\xi \epsilon \mathfrak{X})$$

if and only if $\mathfrak{P}$ is weakly complete relative to $\mathfrak{D}$. Furthermore $\mathfrak{P}$ is weakly complete relative to $\mathfrak{D}$ if and only if it is weakly complete.

[11] R is compositive in case, for every l_1, l_2 in $\mathfrak{L}$, there is an l_3 in $\mathfrak{L}$ such that $l_3 \text{R} l_1$ and $l_3 \text{R} l_2$. (See Moore-Smith, loc. cit., p. 103.)

[12] *Opérations Linéaires*, p. 240.

If $\mathfrak{P}$ is weakly complete relative to $\mathfrak{D}$, then it follows from Theorem III that there exists, for every linear continuous functional M on $\mathfrak{X}$ to $\mathfrak{R}$, a p_0 in $\mathfrak{P}$ with the desired property. Moreover, the existent p_0 is clearly unique. Later it will be shown that the transformation $M_p(\xi) = \xi(p)$ establishes an *equivalence in the sense of Banach*[13] between $\mathfrak{P}$ and $\mathfrak{M}$, the adjoint of $\mathfrak{X}$.

Suppose that every M in $\mathfrak{M}$ is expressible in the form $\xi(p)$ and that p_l is a function on $\mathfrak{L}$ to $\mathfrak{P}$ such that the upper limit as to l of $\| p_l \|$ is finite and the limit as to l of $\xi(p_l)$ exists for each ξ in $\mathfrak{X}$. If $M_l(\xi)$ is defined to be $\xi(p_l)$, there is a linear functional $M_0(\xi)$ which is the limit of $M_l(\xi)$ for each ξ. Furthermore, this limit is continuous since

$$| M_0(\xi) | = \lim_l | M_l(\xi) | \leqq \| \xi \| \cdot \overline{\lim_l} \| p_l \|.$$

Therefore there must exist a p_0 such that $\lim_l \xi(p_l) = M_0(\xi) = \xi(p_0)$.

If one observes that $\mathfrak{P}$ is weakly complete relative to $\mathfrak{D}$ whenever it is weakly complete, then the proof of the theorem is completed.

Corollary 2.1. *If $\mathfrak{X}$ is separable and $\mathfrak{P}$ is complete, then a necessary and sufficient condition that $\mathfrak{P}$ be weakly complete is that it be weakly complete relative to the set of positive integers.*

Clearly it suffices to show that $\mathfrak{P}$ is weakly complete if it is weakly complete relative to the positive integers. For this purpose let $\mathfrak{N}$ be the set of all $M(\xi)$ which are of the form $\xi(p)$. Then $\mathfrak{N}$ is a linear and *total* subset of $\mathfrak{M}$; i.e., if ξ_0 has the property that $N(\xi_0) = 0$ for every N in $\mathfrak{N}$, ξ_0 is the zero functional. Furthermore $\mathfrak{N}$ is weakly closed. For if the sequence $N_n(\xi) = \xi(p_n)$ of functionals in $\mathfrak{N}$ converges weakly to $M(\xi)$, then

$$\lim_n \xi(p_n) = M(\xi) \qquad (\xi \in \mathfrak{X}).$$

But by hypothesis $\mathfrak{P}$ is weakly complete relative to the set of positive integers. Hence there is a p_0 such that

$$\xi(p_0) = \lim_n \xi(p_n) = M(\xi),$$

which proves that M is in $\mathfrak{N}$. Therefore since $\mathfrak{X}$ is separable, $\mathfrak{N}$ coincides with $\mathfrak{M}$.[14] The desired result then follows at once from Theorem IV.

In his treatise on linear operations Banach establishes a sufficient condition that a space be equivalent to the adjoint of its adjoint.[15] The following theorem is a distinct generalization of his result.

Theorem V. *If $\mathfrak{P}$ is weakly complete, then it is equivalent in the sense of Banach to $\mathfrak{M}$ the adjoint of $\mathfrak{X}$.*

[13] Ibid., p. 180; i.e., there is a bi-continuous linear transformation $p = T(M)$ between both spaces such that $\| M \| = \| T(M) \|$.

[14] Ibid., p. 126.

[15] Ibid., p. 189.

The theorem follows from Theorem IV and the remark on equivalence made in the proof of that theorem. To establish this fact, note that the transformation effecting the isomorphism is $M_p(\xi) = \xi(p)$. Therefore $| M_p(\xi) | \leqq \| p \| \cdot \| \xi \|$, and hence the norm of M_p is not greater than the norm of p. But by a theorem of Banach,[16] there is a functional ξ_0 in $\mathfrak{X}$ such that $\xi_0(p) = \| p \|$ and $\| \xi_0 \| = 1$. Hence the norm of M_p equals the norm of p.

The theorem as proved by Banach assumes that $\mathfrak{P}$ is separable, complete, and that every bounded sequence in $\mathfrak{P}$ contains a subsequence which converges weakly to an element of $\mathfrak{P}$. Without assuming either the completeness or separability of $\mathfrak{P}$ it is easy to show that this latter compactness property always implies the weak completeness of $\mathfrak{P}$ relative to the set of positive integers.

3. **The converse of Theorem V.** It is of considerable interest to know when the converse of Theorem V is true. We proceed to investigate that question in this section. Furthermore, we determine the analytic character of any equivalence between $\mathfrak{P}$ and the adjoint of $\mathfrak{X}$.

If φ is a transformation of $\mathfrak{P}$ into $\mathfrak{M}$, then it is clear that $\varphi(p)$ is a function of ξ. For convenience we shall use the notation $\varphi(p \mid \xi_0)$ to denote the functional $M = \varphi(p)$ evaluated at $\xi = \xi_0$.

THEOREM VI. *Let $\mathfrak{P}$ be complete and let φ be a transformation between $\mathfrak{P}$ and the adjoint of $\mathfrak{X}$, which establishes an equivalence in the sense of Banach between these spaces. Then there is a unique rotation*[17] *$f(p)$ of $\mathfrak{P}$ about O_p such that*

$$M(\xi) = \varphi(p \mid \xi) = \xi[f(p)] \qquad (p \in \mathfrak{P}, \xi \in \mathfrak{X}).$$

Since the function $\xi(p)$, as ξ varies over $\mathfrak{X}$, is in $\mathfrak{M}$ and since the vanishing of $\varphi(p \mid \xi)$ for every ξ implies that $p = O_p$, it is clear that there exists a unique function $\rho(p)$ such that

$$\varphi[\rho(p) \mid \xi] = \xi(p) \qquad (p \in \mathfrak{P}, \xi \in \mathfrak{X}). \tag{3.1}$$

Let $\mathfrak{P}_0 \subset \mathfrak{P}$ be the contradomain of ρ, i.e., $\mathfrak{P}_0 = \rho(\mathfrak{P})$. Then ρ is linear and continuous on $\mathfrak{P}$ to $\mathfrak{P}_0$. In fact, since φ establishes an equivalence between $\mathfrak{P}$ and the adjoint of $\mathfrak{X}$ and since the norm of $\xi(p)$, as ξ varies over $\mathfrak{X}$, is the norm of p, we know that

$$\| \rho(p) \| = \| p \|. \tag{3.2}$$

Hence ρ has a linear continuous inverse[18] ρ^{-1} on $\mathfrak{P}_0$ to $\mathfrak{P}$.

It is apparent from what has been said above that $\mathfrak{P}_0$ is a linear closed subset of $\mathfrak{P}$. Therefore it is a Banach space and is of the second category.[19] But $\mathfrak{P}_0$, being the contradomain of a linear continuous operator, is either of the

[16] Ibid., p. 55.
[17] Ibid., p. 173.
[18] Ibid., p. 145.
[19] Ibid., p. 14.

first category or is identical with $\mathfrak{P}$.[20] Therefore $\mathfrak{P}_0 = \mathfrak{P}$, and equation (3.2) implies that ρ is a rotation of $\mathfrak{P}$ about O_p. Hence by equation (3.1) we have

$$\varphi(p \mid \xi) = \xi[f(p)] \qquad (\xi \epsilon \mathfrak{X}),$$

for every p in $\mathfrak{P}$, where f is defined to be ρ^{-1}. Then since ρ is a rotation about O_p, so is f.

THEOREM VII. *If $\mathfrak{P}$ is complete, a necessary and sufficient condition that $\mathfrak{P}$ be equivalent to the adjoint of $\mathfrak{X}$ is that $\mathfrak{P}$ be weakly complete.*

The sufficiency of the condition was shown in Theorem V without assuming the completeness of $\mathfrak{P}$. The necessity will be shown to follow from Theorem VI. Let φ establish an equivalence between $\mathfrak{P}$ and $\mathfrak{M}$. Then $\varphi(p \mid \xi)$ is of the form $\xi[f(p)]$. Suppose that p_δ on $\mathfrak{D}$ to $\mathfrak{P}$ is such that the upper limit with respect to δ of $\| p_\delta \|$ is finite and the limit with respect to δ of $\xi(p_\delta)$ exists for every ξ in $\mathfrak{X}$. It is apparent that there is a functional M on $\mathfrak{X}$ to $\mathfrak{R}$ such that

$$M(\xi) = \lim_{\delta} \varphi[f^{-1}(p_\delta) \mid \xi] \qquad (\xi \epsilon \mathfrak{X}).$$

It follows readily that M is linear and continuous. Therefore, there is a p_0 in $\mathfrak{P}$ such that

$$\varphi[f^{-1}(p_0) \mid \xi] = \lim_{\delta} \varphi[f^{-1}(p_\delta) \mid \xi] \qquad (\xi \epsilon \mathfrak{X}). \tag{3.3}$$

Put into other terms, equation (3.3) states that

$$\xi(p_0) = \lim_{\delta} \xi(p_\delta).$$

Hence, by Corollary 2.1, $\mathfrak{P}$ is weakly complete.

UNIVERSITY OF CHICAGO.

[20] Ibid., p. 38.

SOME EXISTENCE THEOREMS FOR PROBLEMS IN THE CALCULUS OF VARIATIONS

By E. J. McShane

Introduction. In a recent paper[1] I have established a theorem on semi-continuity of integrals of the calculus of variations under hypotheses weak enough to apply to the parametric and ordinary problems, as well as to several other problems. Here I wish to establish existence theorems of a comparable generality for the parametric and ordinary problems, as well as for problems involving higher derivatives.

The added generality in the parametric problem is not very important. It consists merely of a relaxing of the continuity requirements on the integrand; instead of being required to have certain partial derivatives, the integrand is required only to be a lower semi-continuous function of its arguments. The chief point of interest is that the existence theorem for the parametric problem is obtained without added effort as a special case of one of the auxiliary theorems designed to handle the problem in ordinary form.

With problems in ordinary form the situation is quite different. The existence theorems for such problems may be roughly classified into two types: those which depend on the behavior of the integrand $f(x, y, y')$ as $|y'| \to \infty$, and those which depend on the differential properties of minimizing curves. The second type is not considered in this paper. In the first type, a fundamental theorem is the one[2] which applies to quasi-regular integrands for which

$$(*) \qquad f(x, y, y')/|y'| \to \infty \text{ as } |y'| \to \infty.$$

This requirement implies in particular that $\int f dx$ is "positive quasi-regular semi-normal", in Tonelli's terminology. Clearly the requirement that (*) hold everywhere can be relaxed in two ways. We may suppose that (*) fails to hold on a set E, but $\int f dx$ remains positive quasi-regular semi-normal on E. The question is, how general can the set E be without disturbing the existence of the solution? In this paper a class of sets (progressively distributed sets) is defined, and it is shown that E may be any progressively distributed set. All previously known classes of sets E are contained in this class. A different way of relaxing

Received October 28, 1937.

[1] *Semi-continuity of integrals in the calculus of variations*, this Journal, vol. 2 (1936), p. 597. This paper will henceforth be referred to as SC.

[2] M. Nagumo, *Über die gleichmässige Summierbarkeit und ihre Anwendung auf ein Variationsproblem*, Japanese Journal of Mathematics, vol. 6 (1929), pp. 173–182.

E. J. McShane, *Existence theorems for ordinary problems of the calculus of variations*, Annali della R. Sc. Norm. Sup. di Pisa, ser. II, vol. 3 (1934), p. 298.

L. Tonelli, *Su gli integrali del calcolo delle variazioni in forma ordinaria*, Annali della R. Sc. Norm. Sup. di Pisa, ser. II, vol. 3 (1934), p. 400.

(*) is simply to omit it without substituting the hypothesis of semi-normality. The sets E on which this may be done are of two types; those lying on absolutely continuous curves $y = \phi(x)$, and those whose projections on the x-axis have measure 0. The first type[3] we shall not discuss. The second type is here shown to be in a sense somewhat illusory; for under the other hypotheses on $f(x, y, y')$ it is shown that it is possible to change independent variable from x to $\xi = \Lambda(x)$ in such a way that the transform of f satisfies (*). However, it should be remarked that it is only because of the weak continuity requirements on our integrands that this transformation can be used, for the transform of f is not continuous.

This last remark exemplifies the unification of theorems resulting from the analytical generality of the present methods. A previous example was the obtaining of a theorem for problems in parametric form as a by-product of our study of ordinary problems. A third example is contained in §10, where a theorem on integrals involving higher derivatives[4] is deduced without difficulty from the theorem on problems in ordinary form, which latter theorem is in turn immediately deducible from the theorem of §10.

1. **Notation and definitions.** The letters y, η (with or without affixes) will be used to denote vectors $(y^1, \cdots, y^q)$, $(\eta^1, \cdots, \eta^q)$ in q-dimensional space; their lengths will be denoted by $|y|$, $|\eta|$. We shall use a modification of the tensor summation convention: if a *Greek-letter* affix is repeated, the expression is to be summed over the values $1, \cdots, q$ of that affix. Thus $a_{\alpha,n} y_n^\alpha = a_{1,n} y_n^1 + \cdots + a_{q,n} y_n^q$, summed on α but not on n.

For functions $\phi(t)$ the symbol $\dot\phi(t)$ shall denote $\phi'(t)$ whenever $\phi'(t)$ is defined and finite, and shall have the value 0 elsewhere.

Functions shall be permitted to assume the value $+\infty$, but not $-\infty$. For the symbol ∞ we adopt the same (obvious) calculation rules as in SC, and the same (trivial) extension of the meaning of lower semi-continuity. We again use the letters a.c. as an abbreviation for "absolutely continuous", and l.s.c. for "lower semi-continuous". The integrals used will be Lebesgue integrals, with the same minor modification as in SC.

2. The principal objects of study in this paper will be integrals

$$I[y] \equiv \int f(x, y, \dot y)\, dx$$

in which the integrands $f(x, y, \dot y) \equiv f(x, y^1, \cdots, y^q, \dot y^1, \cdots, \dot y^q)$ satisfy the following conditions:

(2.1a) *$f(x, y, \dot y)$ is defined and finite for all (x, y) in a closed set A and all (finite) $\dot y$, and is l.s.c. as a function of its $2q + 1$ arguments;*
(2.1b) *$f(x, y, 0)$ is bounded on A;*
(2.1c) *for each $(x, y) \in A$ the function $f(x, y, \dot y)$ is a convex function of $\dot y$.*

[3] Tonelli, loc. cit., p. 428.
[4] This theorem generalizes the theorem obtained by Cinquini, *Sopra l'esistenza della soluzione*, etc., Annali della R. Sc. Norm. Sup. di Pisa, ser. II, vol. 5 (1936), pp. 169–190.

As is well known, if f has partial derivatives with respect to the $\dot{y}^i$, (2.1c) is satisfied if and only if

$$E(x_0, y_0, \dot{y}_0, \dot{y}) \equiv f(x_0, y_0, \dot{y}) - f(x_0, y_0, \dot{y}_0) - (\dot{y}^\alpha - \dot{y}_0^\alpha) f_{\dot{y}^\alpha}(x_0, y_0, \dot{y}_0) \geqq 0$$

for all $(x_0, y_0) \in A$ and all $\dot{y}_0$ and $\dot{y}$.

From $f(x, y, \dot{y})$ we form the associated parametric integrand $g(x, y, \dot{x}, \dot{y})$, defined by the equations:

$$\text{(2.2)} \qquad \begin{aligned} g(x, y, \dot{x}, \dot{y}) &= \dot{x} f(x, y, \dot{y}/\dot{x}) \quad \text{if } (x, y) \in A \text{ and } \dot{x} > 0, \\ g(x, y, 0, \dot{y}) &= \lim_{\xi \to 0} g(x, y, \xi, \dot{y}) \quad \text{if } (x, y) \in A. \end{aligned}$$

If $f(x, y, \dot{y})$ satisfies conditions (2.1), then $g(x, y, \dot{y})$ satisfies the following conditions:[5]

(2.3a) *g is defined (finite or $+\infty$) for $(x, y) \in A$, $\dot{x} \geqq 0$, all $\dot{y}$, and is finite for $\dot{x} > 0$;*
(2.3b) *g is l.s.c. as a function of all its arguments;*
(2.3c) *for each $(x_0, y_0) \in A$, each $(\dot{x}_0, \dot{y}_0)$ with $\dot{x}_0 \geqq 0$ and each $u < g(x_0, y_0, \dot{x}_0, \dot{y}_0)$, there is a linear function $a\dot{x} + b_\alpha \dot{y}^\alpha$ such that*

$$\text{(i) } a\dot{x}_0 + b_\alpha \dot{y}_0^\alpha > u \qquad \textit{and} \qquad \text{(ii) } g(x_0, y_0, \dot{x}, \dot{y}) \geqq a\dot{x} + b_\alpha \dot{y}^\alpha$$

for all $(\dot{x}, \dot{y})$ with $\dot{x} \geqq 0$;

$$\text{(2.3d)} \qquad g(x, y, t\dot{x}, t\dot{y}) = tg(x, y, \dot{x}, \dot{y}) \quad \textit{if } (x, y) \in A \textit{ and } t \geqq 0.$$

It follows[6] that if C: $x = x(t)$, $y = y(t)$, $(a \leqq t \leqq b)$ is an a.c. representation of a rectifiable curve lying in A, the integral

$$J[C] = \int_a^b g(x(t), y(t), \dot{x}(t), \dot{y}(t))\, dt$$

is defined and is independent of the particular (a.c.) representation of C. In particular, if C happens to have an a.c. representation of the form $y = y(x)$ $(a \leqq x \leqq b)$, we may use this representation to evaluate $J[C]$; then, recalling (2.2), we see that

$$I[y] \equiv \int_a^b f(x, y(x), \dot{y}(x))\, dx = J[C].$$

An important class of integrands f consists of those which, in addition to (2.1), satisfy (for some or all points $(x, y) \in A$) the following condition:

(2.4) *the graph of $f(x, y, \dot{y})$, as a function of $\dot{y}$, has a plane of support which touches it at only a single point;*[7] *that is, there is a linear function $a + b_\alpha \dot{y}^\alpha$ and a particular $\dot{y}_0$ such that*

$$\text{(2.5)} \qquad \begin{aligned} f(x, y, \dot{y}) - a - b_\alpha \dot{y}^\alpha &> 0 \quad \textit{if } \dot{y} \neq \dot{y}_0, \\ f(x, y, \dot{y}_0) - a - b_\alpha \dot{y}_0^\alpha &= 0. \end{aligned}$$

[5] The proof of this statement is contained in SC, p. 613.
[6] SC, Theorem 2.1. Here g satisfies conditions (2.2) of SC.
[7] In the theory of convex bodies such a point is called an *extreme* point.

In case f has partial derivatives with respect to the $\dot{y}^i$, the plane of support at $\dot{y}_0$ is the same as the tangent plane at $\dot{y}_0$, which is the graph of the linear function

$$a + b_\alpha \dot{y}^\alpha \equiv f(x, y, \dot{y}_0) + (\dot{y}^\alpha - \dot{y}_0^\alpha) f_{\dot{y}^\alpha}(x, y, \dot{y}_0);$$

so, if we recall the definition of the E-function, condition (2.3) requires

$$E(x, y, \dot{y}_0, \dot{y}) > 0 \qquad \text{if } \dot{y} \neq \dot{y}_0 .$$

(The other condition, $E(x, y, \dot{y}_0, \dot{y}_0) = 0$, is always satisfied.) The integrals for which f satisfies (2.4) for every $(x, y) \in A$ have been called "positive quasi-regular semi-normal".

We now wish to prove

LEMMA 1. *If $f(x, y, \dot{y})$ satisfies (2.4), then $g(x, y, \dot{x}, \dot{y})$ satisfies the following:*

(2.6) *there is a linear function $c\dot{x} + d_\alpha \dot{y}^\alpha$ such that for all $(\dot{x}, \dot{y}) \neq (0, 0)$ having $\dot{x} \geqq 0$ the inequality*

$$g(x, y, \dot{x}, \dot{y}) - c\dot{x} - d_\alpha \dot{y}^\alpha > 0$$

holds.

With the a and b_i of (2.5) we define $c = a - 1$, $d_i = b_i$. Suppose now $g(x, y, \dot{x}, \dot{y}) - c\dot{x} - d_\alpha \dot{y}^\alpha \leqq 0$ for some $(\dot{x}, \dot{y})$; we shall show that $(\dot{x}, \dot{y}) = (0, 0)$. If we use (2.2) and (2.3),

$$g(x, y, 1, \dot{y}_0) = a + b_\alpha \dot{y}_0^\alpha,$$

$$g(x, y, \dot{x}, \dot{y}) \leqq a\dot{x} - \dot{x} + b_\alpha \dot{y}^\alpha.$$

Since g is convex in $(\dot{x}, \dot{y})$,

$$g(x, y, \tfrac{1}{2}(1 + \dot{x}), \tfrac{1}{2}(\dot{y}_0 + \dot{y})) \leqq \tfrac{1}{2}a(1 + \dot{x}) - \tfrac{1}{2}\dot{x} + \tfrac{1}{2}b_\alpha(\dot{y}^\alpha + \dot{y}_0^\alpha).$$

Dividing both members by $\frac{1}{2}(1 + \dot{x})$ and recalling that g is homogenous, we see that

$$f(x, y, (\dot{y}_0 + \dot{y})/(1 + \dot{x})) = g(x, y, 1, (\dot{y}_0 + \dot{y})/(1 + \dot{x}))$$
$$\leqq a - \dot{x}/(1 + \dot{x}) + b_\alpha(\dot{y}^\alpha + \dot{y}_0^\alpha)/(1 + \dot{x}).$$

Now this, together with (2.5), implies that $\dot{x}/(1 + \dot{x}) \leqq 0$, whence $\dot{x} = 0$. Making the substitution, we get

$$f(x, y, \dot{y}_0 + \dot{y}) \leqq a + b_\alpha(\dot{y}^\alpha + \dot{y}_0^\alpha),$$

which by (2.5) is possible only if $\dot{y} = 0$. This establishes the lemma.

Condition (2.5) attains usefulness through

LEMMA 2. *If $g(x, y, \dot{x}, \dot{y})$ is l.s.c. and satisfies condition (2.6) at a point (x_0, y_0) of A, then there is a positive constant k and a neighborhood U of (x_0, y_0) such that*

$$g(x, y, \dot{x}, \dot{y}) \geqq k[\dot{x}^2 + \dot{y}^\alpha \dot{y}^\alpha]^{\frac{1}{2}} + c\dot{x} + d_\alpha \dot{y}^\alpha \qquad \textit{if } \dot{x} \geqq 0 \textit{ and } (x, y) \in AU.$$

Suppose this false; we can then find a sequence (x_n, y_n) of points of A tending to (x_0, y_0) and a sequence $(\dot{x}_n, \dot{y}_n)$ such that

$$g(x_n, y_n, \dot{x}_n, \dot{y}_n) < [\dot{x}_n^2 + \dot{y}_n^\alpha \dot{y}_n^\alpha]^{\frac{1}{2}}/n + c\dot{x}_n + d_\alpha \dot{y}_n^\alpha.$$

This implies at once that $(\dot{x}_n, \dot{y}_n) \neq (0, 0)$. As both members are positively homogeneous of degree 1 in $(\dot{x}, \dot{y})$, we may as well suppose that the expression in square brackets on the right has the value 1. From the $(\dot{x}_n, \dot{y}_n)$ we can select a subsequence $(\dot{x}_m, \dot{y}_m)$ converging to a limit $(\dot{x}_0, \dot{y}_0)$; this limit will satisfy $\dot{x}_0^2 + \dot{y}_0^\alpha \dot{y}_0^\alpha = 1$. Now, since g is l.s.c.,

$$g(x_0, y_0, \dot{x}_0, \dot{y}_0) \leqq \liminf_{m\to\infty} g(x_m, y_m, \dot{x}_m, \dot{y}_m) \leqq c\dot{x}_0 + d_\alpha \dot{y}_0^\alpha.$$

This violates condition (2.6), and so Lemma 2 is established.

3. In the paper cited in footnote 1, some theorems on semi-continuity were established and utilized, together with the simplest convergence theorem, to obtain general existence theorems. This simplest convergence theorem applies only to integrals for which $g(x, y, \dot{x}, \dot{y}) > 0$ if $(\dot{x}, \dot{y}) \neq (0, 0)$. In this section we shall establish a more general convergence theorem, of the type already studied by Hahn[8] and Tonelli.[9]

LEMMA 3. *If*

(a) *A is a bounded closed set;*

(b) *$J[C]$ satisfies conditions* (2.3) *and* (2.6) *for all* $(x, y) \in A$;

(c) *for each constant x_0, there is a number $\mu(x_0)$ such that all the curves*

$$C\colon x = x_0, \quad y^i = y^i(t), \qquad (a \leqq t \leqq b)$$

in A for which $J[C] \leqq 0$ have lengths $\leqq \mu(x_0)$;
then for every positive number M all[10] *curves C in A such that $J[C] \leqq M$ have lengths less than a number N depending only on M.*

In the course of the proof of this lemma we shall consider curves C with various affixes. For such curves the defining functions, etc., will have corresponding affixes; thus, if C_n^r be a curve, we understand that it is defined by functions $x = x_n^r(s)$, $y = y_n^r(s)$, $(0 \leqq s \leqq L_n^r)$, where L_n^r is the length of C_n^r. Also, we shall have occasion to select from among a finite number of arcs of a curve one which satisfies a certain criterion; if several of the arcs satisfy the criterion, we shall always choose the one on which s is smallest. Finally, given a curve C_n, we define $J_n(s)$ to be the integral $\int_0^s g(x_n, y_n, \dot{x}_n, \dot{y}_n)\,ds$ $(0 \leqq s \leqq L_n)$.

Suppose now that the lemma is false. We can then find an M and a sequence $\{C_n\}$ (with $\dot{x}_n(s) \geqq 0$) such that

$$J[C_n] \leqq M, \tag{3.1}$$

and

$$\lim L_n = \infty. \tag{3.2}$$

[8] H. Hahn, *Über ein Existenztheorem der Variationsrechnung*, Sitzungber. d. Akad. d. Wiss. Wien, Mathem.-Naturwiss. Kl., Abt. II-a, vol. 134 (1925), pp. 437–447.

[9] L. Tonelli, *Sull'esistenza del minimo in problemi di calcolo delle variazioni*, Annali della R. Sc. Norm. Sup. di Pisa, ser. II, vol. 1 (1932), pp. 89–99.

[10] We continue to assume $\dot{x} \geqq 0$.

We shall first show that the sequence $\{C_n\}$ can be taken so as to have the following properties:

(α) $\liminf J_n(s) \leqq 0$ for all $s \geqq 0$,

(β) $\lim L_n = \infty$,

(γ) the sequences $\{x_n(0)\}$ and $\{x_n(L_n)\}$ converge to a common limit ξ.

In order to prove this we find it convenient to consider two cases.

Case I. The oscillations Osc $J_n(s)$ are unbounded. Let $J_n(s)$ assume its maximum at $s_{n,2}$ and its minimum at $s_{n,1}$. Either $J_n(s_{n,2})$ is unbounded above, in which case we may assume $\lim J_n(s_{n,2}) = \infty$; or else $J_n(s_{n,1})$ is unbounded below, in which case we may assume $\lim J_n(s_{n,1}) = -\infty$. In the former case we select the arc corresponding to $s_{n,2} \leqq s \leqq L_n$ and rename it C_n^*; in the latter case we denote by $s_{n,0}$ the greatest s in the interval $[0, s_{n,1}]$ for which $J_n(s) = 0$, and choose C_n^* to be the arc corresponding to $s_{n,0} \leqq s \leqq s_{n,1}$. In either case we have

$$\lim J[C_n^*] = -\infty. \tag{3.3}$$

We may assume that $J[C_n^*] \leqq -n^2$; if this is not the case, we have only to choose a subsequence of $\{C_n^*\}$, as is shown by (3.3). For $k = 0, 1, 2, \cdots, n$ we define s_k to be the greatest s for which $J_n^*(s) = kn^{-1}J[C_n^*]$. The curve C_n^* is thus cut into n arcs. If d be the diameter of the set A, for at least one such arc the projection on the x-axis is $\leqq d/n$; we choose one such arc and call it $\bar{C}_n$. Then by the definition of the s_k we have

$$\bar{J}_n(s) = J_n^*(s_k + s) - J_n^*(s_k) \leqq 0 \qquad (0 \leqq s \leqq \bar{L}_n), \tag{3.4}$$

$$\int_0^{\bar{L}_n} g(\bar{x}_n, \bar{y}_n, \dot{\bar{x}}_n, \dot{\bar{y}}_n)\, ds = J[\bar{C}_n] = \frac{1}{n} J[C_n^*] \leqq -n.$$

The set of all arguments $(x, y, \dot{x}, \dot{y})$ with $(x, y) \in A$, $\dot{x} \geqq 0$, $\dot{x}^2 + \dot{y}^\alpha \dot{y}^\alpha = 1$ is bounded and closed, and on it g is l.s.c.; hence on this set g attains its lower bound $-\delta$. From (3.4) it is obvious that $-\delta < 0$, and

$$-\delta \bar{L}_n = -\int_0^{\bar{L}_n} \delta \cdot ds \leqq J[\bar{C}_n] \leqq -n,$$

so that

$$\bar{L}_n \geqq n/\delta, \tag{3.5}$$

and

$$\bar{L}_n \to \infty.$$

From the curves $\bar{C}_n$ we select a subsequence C_n for which the numbers $\bar{x}_n(0)$ approach a limit ξ. But the projection of $\bar{C}_n$ on the x-axis is at most d/n, so that $\bar{x}_n(0) \leqq \bar{x}(\bar{L}_n) \leqq \bar{x}_n(0) + d/n$. This implies that $\lim x_n(L_n) = \xi$, and conditions (α), (β), (γ) are all satisfied.

Case II. The oscillations Osc $J_n(s)$ are bounded. Suppose that Osc $J_n(s) \leqq$

K. We assume, as we may, that $L_n \geqq n^3$, and we cut C_n into n arcs each of length L_n/n. At least one such arc has a projection on the x-axis which is $\leqq d/n$, where d is the diameter of A. We choose such an arc and call it C_n^*. Clearly $L_n^* \geqq n^2$ and Osc $J_n^*(s) \leqq K$. If we choose a subsequence (which we again call C_n^*) such that the numbers $x_n^*(0)$ converge to a limit ξ, then from the inequality $x_n^*(0) \leqq x_n^*(L_n^*) \leqq x_n^*(0) + d/n$ we see that $\lim x_n^*(L_n^*) = \xi$. Hence (γ) is satisfied.

If for infinitely many curves C_n^* we have $J_n^*(s) \leqq 0$ $(0 \leqq s \leqq L_n^*)$, we select these curves as the C_n, and (α), (β), (γ) are satisfied. If not, we have for almost all n the inequality

$$m_n = \max J_n^*(s) > 0.$$

For $k = 0, 1, \cdots, n$ we define s_k to be the least s for which $J_n^*(s) = m_n k/n$. The curve C_n^* is thus divided into n arcs. At least one has length $\geqq L_n^*/n \geqq n$; we choose one such arc and call it C_n. Since the C_n are arcs of the C_n^*, condition (γ) remains satisfied; and since $L_n \geqq n$, (β) is also satisfied. By the definition of the s_k we have (for a certain k)

$$J_n(s) = J_n^*(s_k + s) - J_n^*(s_k) \leqq m_n(k+1)/n - m_n k/n \leqq K/n \; (0 \leqq s \leqq L_n);$$

and thus (α) also holds. So in this case also conditions (α), (β) and (γ) are satisfied.

Letting h be a positive integer, we consider the functions $x = x_n(s)$, $y = y_n(s)$, $(0 \leqq s \leqq h)$, which are defined on this interval for almost all n. Since these functions are uniformly bounded and satisfy a Lipschitz condition of constant 1, we can select a subsequence $\{C_{n,h}\}$ for which $x_{n,h}(s)$ and $y_{n,h}(s)$ converge to limit functions $x_0(s)$, $y_0(s)$, respectively, uniformly for $0 \leqq s \leqq h$. If we thus select $\{C_{n,1}\}$, and from it select $\{C_{n,2}\}$, and so forth, we can by the diagonal process select a subsequence $\{C_m\}$ such that on every finite interval $0 \leqq s \leqq s_0$ the functions $x_m(s)$, $y_m(s)$ converge uniformly to limit functions $x_0(s)$, $y_0(s)$. Because of (γ), $x_0(s) \equiv \xi$. Because of the lower semi-continuity[11] of $J[C]$, it follows from (α) that $J_0(s) \leqq 0$ for every s. Hence by hypothesis (c) the curve $x = x_0(s)$, $y = y_0(s)$, $(0 \leqq s \leqq s_0)$ has length $\leqq \mu(\xi)$, whatever s_0 may be. This can only be true if $x_0(s)$, $y_0(s)$ approach unique limits x^*, y^* as $s \to \infty$. (Clearly $x^* = \xi$.)

By Lemmas 1 and 2, hypothesis (b) implies that there exists a neighborhood U of (x^*, y^*) (which we shall take to be a spherical neighborhood having radius 2ρ and center at (x^*, y^*)) and numbers $k > 0$, a, $b^1, \cdots, b^q$ such that

$$g(x, y, \dot{x}, \dot{y}) - a\dot{x} - b^\alpha \dot{y}^\alpha \geqq k$$

for all points (x, y) of AU and all unit vectors $(\dot{x}, \dot{y})$ with $\dot{x} \geqq 0$. Let N be chosen large enough so that $(x_0(s), y_0(s))$ has distance less than ρ from (x^*, y^*) when $s \geqq N$, and let h be any positive number. Then for all sufficiently large n

[11] SC, Theorem 3.1.

the arc $x = x_n(s)$, $y = y_n(s)$, $(N \leqq s \leqq N + h)$ lies in U. Since $(\dot{x}_n, \dot{y}_n)$ is a unit vector for almost all s, and $\dot{x}_n \geqq 0$, this implies

$$\int_N^{N+h} g(x_n, y_n, \dot{x}_n, \dot{y}_n)ds \geqq kh + a[x_n(N+h) - x_n(N)] + b^\alpha[y_n^\alpha(N+h) - y_n^\alpha(N)]$$
$$\geqq kh - 4\rho(|a| + |b^1| + \cdots + |b^q|).$$

Using again the semi-continuity of $J[C]$, we find that for all sufficiently large n we have

$$\int_0^N g(x_n, y_n, \dot{x}_n, \dot{y}_n)\,ds \geqq J_0(N) - 1.$$

Adding the last two inequalities, we get

$$J_n(N + h) \geqq kh + J_0(N) - 1 - 4\rho(|a| + |b^1| + \cdots + |b^q|).$$

If we choose and fix a sufficiently large value of h, the right side exceeds 1. But this contradicts (α), and so the lemma is established.

4. We are now able to establish an existence theorem for the associated parametric problem $\int g ds = \min$.

THEOREM 1. *If*

(a) *A is bounded and closed, and $g(x, y, \dot{x}, \dot{y})$ satisfies conditions* (2.3) *on A*;

(b) *g satisfies condition* (2.6) *for each $(x, y) \in A$*;

(c) *for each fixed x_0, all the rectifiable curves C: $x = x_0$, $y = y(s)$ with $J[C] \leqq 0$ have uniformly bounded lengths*;

(d) *$\bar{K}$ is a complete class of rectifiable curves lying in A and having $\dot{x}(s) \geqq 0$; then in the class $\bar{K}$ there is a minimizing curve for $J[C]$.*

If $J[C] = \infty$ for every curve of $\bar{K}$, then any curve of $\bar{K}$ gives $J[C]$ its least value, ∞. Otherwise $J[C]$ assumes a finite value $M - 1$. Choose now a sequence $\{C_n\}$ of curves of $\bar{K}$ such that $J[C_n]$ tends to the lower bound m of $J[C]$ on $\bar{K}$. We may assume $J[C_n] < M$ for all n. By Lemma 3, the curves C_n have uniformly bounded lengths. Therefore it is possible to select a subsequence $\{C_{n_i}\}$ of the C_n such that C_{n_i} tends to a rectifiable limit curve C_0. Since $\bar{K}$ is complete, C_0 is in $\bar{K}$, and so $J[C_0] \geqq m$. But on the other hand the curves $C_0, C_{n_1}, C_{n_2}, \cdots$ are of uniformly bounded lengths, so $J[C]$ is l.s.c. on this class,[12] and $J[C_0] \leqq \liminf J[C_{n_i}] = m$. Therefore $J[C_0] = m$, and this completes the proof.

From this theorem we obtain as a corollary an existence theorem for integrals in parametric form. An integrand $g(y, \dot{y})$ in parametric form may be regarded as an integrand $g(x, y, \dot{x}, \dot{y})$ which happens to be independent of x and $\dot{x}$. If $g(y, \dot{y})$ is defined for $y \in B$ and all y, then we may regard it as a function $g(x, y, \dot{x}, \dot{y})$ defined for all x, all $y \in B$, all $\dot{x} \geqq 0$, all $\dot{y}$. Suppose that $g(y, \dot{y})$ satisfies the condition

[12] SC, Theorem 3.1.

(4.1) *for each* $y \in B$, *each* $\dot{y}_0$ *and each* $u < g(y, \dot{y}_0)$ *there is a linear function* $c_\alpha \dot{y}^\alpha$ *such that*

(i) $$c_\alpha \dot{y}_0^\alpha > u$$

and

(ii) $$g(y, \dot{y}) \geqq c_\alpha \dot{y}^\alpha$$

for all $\dot{y}$.

Then clearly condition (2.3c) is satisfied if we take $a = 0$, $b_i = c_i$. In particular, condition (4.1) (and therefore (2.3c)) holds if $g(y, \dot{y})$ is convex in $\dot{y}$ for each $y \in B$. Furthermore, if

(4.2) *for each* $y \in B$ *there is a linear function* $c_\alpha \dot{y}^\alpha$ *such that*

$$g(y, \dot{y}) - c_\alpha \dot{y}^\alpha > 0 \quad \text{if } | \dot{y} | > 0,$$

then condition (2.6) holds; we have only to take $d_i = c_i$, $c = -1$ in (2.6). We may therefore state

THEOREM 2. *If*

(a) *$g(y, y)$ is defined (finite or $+\infty$) and l.s.c. for all y in a bounded closed set B and all y, and is positively homogeneous of degree* 1 *in y;*

(b) *g satisfies conditions* (4.1) *and* (4.2);

(c) *the rectifiable curves C: $y = y(s)$ in B for which $J[C] \leqq 0$ have uniformly bounded lengths;*

(d) *K is a complete class of rectifiable curves lying in B;*

then in the class K there is a minimizing curve for $J[C]$.

In (x, y)-space we consider the set A of points (x, y) such that $x = 0$ and $y \in B$; this is bounded and closed. We set $g(x, y, \dot{x}, \dot{y}) = g(y, \dot{y})$ for $(x, y) \in A$, and let $\bar{K}$ be the class of curves C: $x = 0$, $y = y(s)$ with $y(s)$ in K. Then by Theorem 1 there is a minimizing curve $x = 0$, $y = y_0(s)$ for $\int g(x, y, \dot{x}, \dot{y})\, ds$ in the class $\bar{K}$. It follows at once that $y = y_0(s)$ minimizes $J[C]$ in K.

5. The remainder of this paper will be primarily concerned with existence theorems for integrals in ordinary form. For the study of these integrals it is convenient to utilize a property of point sets which we shall call "progressive distribution".

Roughly speaking, we may say that a set E of points (x, y) has this property if a rectifiable curve $x = x(s)$, $y = y(s)$ with $x'(s) \geqq 0$ cannot stay in the set E without obliging the tangent vector (x', y') to have a positive component x' along the x-axis. More precisely:

(5.1) *A set E in $(x, y^1, \cdots, y^q)$-space is progressively distributed if for every absolutely continuous representation of every rectifiable curve $x = x(t)$, $y = y(t)$, $(a \leqq t \leqq b)$, $\dot{x}(t) \geqq 0$ the conditions $(x(t), y(t))$ in E, $\dot{x}(t) = 0$, $\dot{y}(t) \neq 0$ are not simultaneously satisfied except at most on a set of t of measure* 0.

If we introduce the notation (to be retained throughout this section) T_E for the set of t such that $(x(t), y(t)) \in E$, T_0 for the set of t such that $\dot{x}(t) = 0$, and T_+ for the set of t such that $| \dot{y}(t) | > 0$, condition (5.1) assumes the form

$$m(T_0 T_+ T_E) = 0.$$

An equivalent statement is that

(5.2) *a set E is progressively distributed if, whenever $x = x(s)$, $y = y(s)$ is a rectifiable curve (s being arc length) with $\dot{x}(s) \geqq 0$, the conditions $\dot{x}(s) = 0$ and $(x(s), y(s)) \in E$ cannot hold simultaneously except on a set of measure* 0.

With symbols analogous to those introduced just above, condition (5.2) can be written

$$m(S_0 S_E) = 0.$$

It is clear that (5.1) implies (5.2). For suppose (5.1) holds; we must show $m(S_0S_E) = 0$. Write

$$S_0S_E = S_0S_ES_+ + S_0S_ECS_+ .$$

The first set on the right has measure 0 by (5.1). The second set is contained in the set S_0CS_+ on which $\dot{x}(s) = 0$ and $|\dot{y}(s)| = 0$. This set has measure 0, so S_0S_E has measure 0.

Conversely, let the set E satisfy (5.2). Let $x = x(s)$, $y = y(s)$, $(0 \leqq s \leqq L)$ be a rectifiable curve with arc length as parameter and such that $\dot{x}(s) \geqq 0$, and let $x = \xi(t)$, $y = \eta(t)$, $(a \leqq t \leqq b)$ be another a.c. representation of the same curve. Define N_s to be the set of s such that $x'(s)$ or $y'(s)$ is undefined, and define N_t to be the set of t for which one or more of the derivatives $s'(t)$, $\xi'(t)$, $\eta'(t)$ is undefined. Since all these functions are a.c., $mN_s = mN_t = 0$.

From the inequality

$$|\eta^i(t_1) - \eta^i(t_2)| \leqq |s(t_1) - s(t_2)|$$

we see at once that $\eta'(t) = 0$ whenever $s'(t) = 0$. If $t \in T_ET_0T_+ - N_t$, then $\eta'(t)$ is defined and $\neq 0$, so $s'(t) \neq 0$:

$$(5.3) \qquad s'(t) \neq 0 \quad \text{for } t \in T_ET_0T_+ - N_t .$$

We next show

$$(5.4) \qquad s(t) \in S_0S_E + N_s \quad \text{for } t \in T_ET_0T_+ - N_t .$$

If t is in the last-mentioned set, either $x'(s(t))$ and $y'(s(t))$ are both defined or else they are not. In the latter case $s(t) \in N_s$ by definition of N_s. In the former case, $\xi'(t) = x'(s(t))\, s'(t)$. But $\xi'(t) = 0$ because $t \in T_0$, whereas $s'(t) \neq 0$ by (5.3). Hence $x'(s(t)) = 0$, and $s(t) \in S_0$. That $s(t) \in S_E$ is immediate, because $(x(s(t)), y(s(t))) = (\xi(t), \eta(t)) \in E$. So in this case $s(t) \in S_0S_E$, and (5.4) is established.

Now S_0S_E has measure 0 by condition (5.2), while we already know N_s has measure 0. So by (5.4), the a.c. function $s(t)$ maps $T_ET_0T_+ - N_t$ on a set of measure 0, and therefore[13] $\dot{s}(t) = 0$ almost everywhere in $T_0T_ET_+ - N_t$. This, with (5.3), implies that the last-named set has measure 0. But N_t has measure 0, so

$$m(T_0T_ET_+) = 0,$$

and condition (5.1) is satisfied.

[13] Carathéodory, *Vorlesungen über Reelle Funktionen*, p. 560, Theorem 2.

We shall now investigate some of the properties of progressively distributed sets.

LEMMA 4. *If E is progressively distributed, so is every subset E^* of E.*

The proof is obvious.

LEMMA 5. *If $E_1, E_2, \cdots$ is a finite or denumerable collection of progressively distributed sets, the set $E = \sum E_n$ is progressively distributed.*

With the above notation we have

$$T_E = \sum T_{E_n}, \qquad T_0T_+T_E = \sum T_0T_+T_{E_n},$$

and each set in the sum on the right has measure 0.

LEMMA 6. *If the set E is such that its projections on the spaces $(x, y^1, \cdots, y^k)$ and $(x, y^{k+1}, \cdots, y^q)$ are progressively distributed, then E is progressively distributed.*

Let the projections mentioned be E', E'', respectively. With the usual notation, the set T_+ is the sum of the sets T'_+ on which $(\dot{y}^1, \cdots, \dot{y}^k) \neq (0, \cdots, 0)$ and the set T''_+ on which $(\dot{y}^{k+1}, \cdots, \dot{y}^q) \neq (0, \cdots, 0)$. The set T_E is the product of the sets T'_E, on which $(x, y^1(x), \cdots, y^q(x)) \in E'$, and T''_E, on which

$$(x, y^{k+1}(x), \cdots, y^q(x)) \in E''.$$

By hypothesis,

$$m(T_0T'_ET'_+) = m(T_0T''_ET''_+) = 0.$$

Hence the set

$$T_0T_ET_+ = T_0T'_ET''_E(T'_+ + T''_+) \subset T_0T'_ET'_+ + T_0T''_ET''_+$$

has measure 0, and E is progressively distributed.

By induction we obtain the corollary

LEMMA 7. *If there is a partition of the y^i into sets $(y^1, \cdots, y^a)$, $(y^{a+1}, \cdots, y^b)$, $\cdots$, $(y^{h+1}, \cdots, y^q)$ such that the projection of E on each of the spaces*

$$(x, y^1, \cdots, y^a), \cdots, (x, y^{h+1}, \cdots, y^q)$$

is progressively distributed, then E is progressively distributed.

LEMMA 8. *If E_1 is progressively distributed and E is the set of all points (x, y) such that $(x, y + \phi(x)) \in E_1$, where $\phi^1(x), \cdots, \phi^q(x)$ are a.c. functions, then E is progressively distributed.*

Let $x = x(t)$, $y = y(t)$, $(a \leqq t \leqq b)$ be an a.c. curve with $\dot{x} \geqq 0$, and let $y_1(t) = y(t) + \phi(x(t))$. Since $\phi(x(t))$ is a.c.,[14] the functions $x = x(t)$, $y = y_1(t)$, $(a \leqq t \leqq b)$ also are an a.c. representation of a curve with $\dot{x} \geqq 0$. In accordance with our usual notation, we let T_E be the set of all t such that $(x(t), y(t)) \in E$, and let $T^1_{E_1}$ be the set of all t such that $(x(t), y_1(t)) \in E_1$. By the definition of y_1, these sets are identical. Also, we let T_+, T^1_+ be respectively the sets of t such that $|\dot{y}(t)| \neq 0$, $|\dot{y}_1(t)| \neq 0$. By hypothesis, $m(T_0T^1_+T^1_{E_1}) = 0$; we must show $m(T_0T_+T_E) = 0$. Since $x(t)$ is a.c. and $\dot{x} = 0$ on T_0, the set T_0 is mapped by

[14] Carathéodory, op. cit., p. 556, Theorem 12.

$x = x(t)$ on a set of x of measure 0.[15] Since $\phi^i(x)$ is a.c., this last set is in turn mapped by $y^i = \phi^i(x)$ on a set Y^i of measure 0. That is, $y^i = \phi^i(x(t))$ is an a.c. function which maps T_0 on a set Y^i of measure 0. Therefore,[16] for almost all $t \in T_0$, we have $d\phi^i/dt = 0$, so that $\dot{y}(t) = \dot{y}_1(t) - d\phi(x(t))/dt = \dot{y}_1(t)$ for almost all $t \in T_0$. Thus the sets T_0T_+ and $T_0T_+^1$ differ only by a set of measure 0, and so

$$m(T_0T_+T_E) = m(T_0T_+^1T_{E_1}^1) = 0.$$

LEMMA 9. *If E_1 is progressively distributed and $\xi(x)$ is monotonic increasing and a.c. on an interval (x_1, x_2), then the set E of all points (x, y) with $x_1 \leqq x \leqq x_2$ and $(\xi(x), y) \in E_1$ is progressively distributed.*

Let $x = x(t)$, $y = y(t)$ be an a.c. curve with $\dot{x} \geqq 0$; then $x = x_1(t) = \xi(x(t))$, $y = y(t)$ is also an a.c. curve with $\dot{x}_1 \geqq 0$. The set T_E on which $(x(t), y(t)) \in E$ is the same as the set $T_{E_1}^1$ on which $(x_1(t), y(t)) \in E_1$. The set T_+ is the same for both curves. If we can show that the set T_0^1, on which $\dot{x}_1 = 0$, and the set T_0, on which $\dot{x} = 0$, are such that $m(T_0 - T_0^1) = 0$, the proof will then be complete, for then we will have

$$m(T_0T_+T_E) \leqq m[(T_0 - T_0^1)T_+T_E] + m(T_0^1T_+T_E) = 0.$$

Let N_x be the set of x such that $\xi'(x)$ is undefined; this set has measure 0, since $\xi(x)$ is a.c. For the same reason, its image N_ξ, consisting of all numbers $\xi(x)$ with $x \in N_x$, has measure 0. We now split T_0 into the subset T' on which $x(t) \in N_x$, and the subset T'', on which $x(t)$ is not in N_x, and $\xi'(x(t))$ is therefore defined. The a.c. function $x_1(t) = \xi(x(t))$ maps T' on a subset of N_ξ, which has measure 0; so[17] we have $\dot{x}_1(t) = 0$ almost everywhere in T'. On almost all of T'' we have $\dot{x}_1(t) = \xi'(x(t))\dot{x}(t) = 0$. Hence $\dot{x}_1(t) = 0$ almost everywhere in T_0; that is, $T_0 - T_0^1$ is a set of measure 0, and the proof is complete.

LEMMA 10. *In the plane, if the projection P of the set E on the y-axis has measure 0, the set E is progressively distributed.*

If $t \in T_E$, then $y(t) \in P$. This implies[17] $\dot{y}(t) = 0$ almost everywhere in T_E; that is, $m(T_+T_E) = 0$, and E is progressively distributed.

By Lemmas 5 to 10 we see, for example, that a set E is progressively distributed if it lies on a family of curves $y^i = \phi_n^i(x) + \alpha^i$, where the $\phi_n^i(x)$ belong to a denumerable family of a.c. functions and the numbers α^i belong to a set of measure 0.

The application of the notion of progressively distributed sets to our variations problems will be through

LEMMA 11. *If $g(x, y, \dot{x}, \dot{y})$ satisfies conditions (2.3) on A, and there is a progressively distributed set E such that $g(x, y, 0, \dot{y}) = \infty$ for all $(x, y) \in A - E$ and all $\dot{y} \neq 0$, then for every a.c. curve C: $x = x(s)$, $y = y(s)$, $(0 \leqq s \leqq L)$ such that $J[C] < \infty$ we have $\dot{x}(s) > 0$ almost everywhere.*

Since $J[C] < \infty$, the function $g(x, y, \dot{x}, \dot{y})$ is summable, hence almost everywhere finite. Split $(0, L)$ into the set S_E, on which $(x(s), y(s)) \in E$, and the

[15] Hobson, *Theory of Functions of a Real Variable*, vol. 1, pp. 606 and 342.

[16] Carathéodory, loc. cit. (footnote 13).

[17] Carathéodory, loc. cit. (footnote 13).

remaining set CS_E. Almost everywhere in S_E we have $\dot{x}(s) > 0$, since E is progressively distributed. Almost everywhere in CS_E we have $g(x, y, \dot{x}, \dot{y}) < \infty$, hence $\dot{x} > 0$. This completes the proof.

Let us agree that a function $y(x)$, defined and single-valued on a degenerate interval consisting of a single point, will be called a.c. Then Lemma 11 implies at once

LEMMA 12. *Under the hypotheses of Lemma* 11, *if C is a rectifiable curve such that $J[C] < \infty$, then C has an a.c. representation $y = y(x)$ $(a \leqq x \leqq b)$.*

Let $x = x(s)$, $y = \bar{y}(s)$, $(0 \leqq s \leqq L)$ be the representation of C with arc length as parameter. If $L = 0$, C consists of a single point, which can be given the a.c. representation $y = y(x) \equiv \bar{y}(0)$, $x = x(0)$. Otherwise, $L > 0$. Then, by Lemma 11, $\dot{x}(s) > 0$ for almost all s, so[18] $x(s)$ has an a.c. inverse $s(x)$. Changing parameter to x, we find that C has the representation $x = x$, $y = y(x) \equiv \bar{y}(s(x))$, $(x(0) \leqq x \leqq x(L))$ and $y(x)$ is a.c.[19]

Another corollary is

LEMMA 13. *Under the hypotheses of Lemma* 11, *hypothesis* (c) *of Lemma* 3 *is fulfilled.*

Let $x = x(s) \equiv x_0$, $y = y(s)$, $(0 \leqq s \leqq L)$ be a curve with $J[C] \leqq 0$. Then $\dot{x}(s) \equiv 0$, while by Lemma 11 $\dot{x}(s) > 0$ for almost all s in $[0, L]$. This is possible only if $L = 0$, so in hypothesis (c) of Lemma 3 we can take $\mu(x_0) = 0$.

6. We now take up the proof of our first existence theorem for problems in ordinary form.

THEOREM 3. *If*

(a) *A is a bounded closed set of points;*

(b) *K_a is a complete class of a.c. functions lying in A;*

(c) *$f(x, y, \dot{y})$ satisfies conditions* (2.1) *and* (2.4) *on A;*

(d) *the equation*[20]

$$\lim_{\xi \to 0} \xi f(x, y, \dot{y}/\xi) = \infty \tag{6.1}$$

holds for all $\dot{y} \neq (0, \cdots, 0)$ and all (x, y) in $A - E$, where E is a progressively distributed set;

then there exists a function $y = y_0(x)$ in the class K_a for which $I[y]$ assumes its least value on K_a.

If $I[y] = \infty$ for all curves y of K, then any curve y_0 of K_a gives $I[y]$ its least

[18] Carathéodory, loc. cit. (footnote 13), p. 584, Satz 3.

[19] Carathéodory, loc. cit. (footnote 14).

[20] It is worth noticing that the associated parametric integrand $g(x, y, \dot{x}, \dot{y})$ necessarily satisfies condition (2.6) at each $(x, y) \in A$ at which (6.1) (or the equivalent statement (6.3)) holds. For let S_n be the set of $(\dot{x}, \dot{y})$ for which $\dot{x}^2 + \dot{y}^\alpha \dot{y}^\alpha = 1$ and $g(x, y, \dot{x}, \dot{y}) + n\dot{x} \leqq 0$. Each S_n is bounded and closed, and $S_1 \supset S_2 \supset S_3 \supset \cdots$. No point $(\dot{x}, \dot{y})$ with $\dot{x} > 0$ belongs to any S_n with $n > -g(x, y, \dot{x}, \dot{y})/\dot{x} \neq +\infty$, and no point $(0, \dot{y})$ belongs to any S_n, since $g(x, y, 0, \dot{y}) = \infty$ by (6.3). So ΠS_n is empty, and therefore there is an n for which S_n is empty. For this n we have $g + n\dot{x} > 0$ if $\dot{x}^2 + \dot{y}^\alpha \dot{y}^\alpha = 1$, and by homogeneity $g + n\dot{x} > 0$ if $(\dot{x}, \dot{y}) \neq (0, 0)$.

value, ∞. Otherwise, for some y of K_a the integral $I[y]$ assumes a finite value $M - 1$. Choose now a sequence $y_1, y_2, \cdots$ of functions of K_a for which $I[y_n]$ tends to the lower bound m of $I[y]$ on K_a. We may assume $I[y_n] < M$ for all n. That is, using the associated parametric integrand and denoting the curve $y = y_n(x)$ by C_n, we have

$$J[C_n] < M, \qquad J[C_n] \to m. \tag{6.2}$$

In terms of the associated parametric integrand, (6.1) becomes

$$g(x, y, 0, \dot{y}) = \infty \quad \text{for } (x, y) \in A - E, \mid \dot{y} \mid \neq 0. \tag{6.3}$$

Therefore the hypotheses of Lemma 11 are satisfied, and by Lemma 13 hypothesis (c) of Lemma 3 is verified. Moreover, from Lemma 1 we learn that g satisfies condition (2.6) for all $(x, y) \in A^*$. So by Lemma 3 the curves C_n have uniformly bounded lengths. Therefore, by Hilbert's theorem, they possess a curve of accumulation $\bar{C}$: $x = \bar{x}(s)$, $y = \bar{y}(s)$, and a subsequence of the C_n (we may suppose it is the whole sequence) tends to $\bar{C}$. By the semi-continuity[21] of $J[C]$ it follows that $J[\bar{C}] \leqq \liminf J[C_n] = m$. But by Lemma 12 this implies that $\bar{C}$ has an a.c. representation $y = y_0(x)$. Since $y = y_0(x)$ is an a.c. limit function of a sequence $\{y_n\}$ of functions of K_a, and K_a is complete, $y = y_0(x)$ is itself in K_a, and $I[y_0] \geqq m$. Therefore,

$$I[y_0] = J[\bar{C}] = m,$$

and the theorem is proved.

7. So far we have assumed that (except on a progressively distributed set) the relation $g(x, y, 0, \dot{y}) = \infty$ holds for $\mid \dot{y} \mid \neq 0$. We now show that our conclusions hold if by a change of independent variable we can arrive at this property for the transform of g. Let us say that

$$f(x, y, \dot{y}) \text{ (or } g(x, y, \dot{x}, \dot{y})) \tag{7.1}$$

satisfies the x-transform condition on A with the exception of E if there is
(7.1a) *a function $\nu(x)$ defined on $[a, b]$ (where a and b are such that the set A lies between $x = a$ and $x = b$), positive, l.s.c. and summable;*
(7.1b) *a set N whose projection on the x-axis has measure* 0, *such that if*

$$(x_n, y_n, \dot{x}_n, \dot{y}_n) \to (x_0, y_0, \dot{x}_0, \dot{y}_0)$$

and

$$(x_n, y_n) \in A - N, \qquad (x_0, y_0) \in A - E, \qquad \mid \dot{y}_0 \mid \neq 0,$$

then

$$\lim_{n \to \infty} g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) = \infty.$$

Then

[21] SC, Theorem 5.1.

THEOREM 4. *If*

(a) *A is a bounded closed set, and $f(x, y, \dot{y})$ satisfies condition* (2.1) *on A;*

(b) *$f(x, y, \dot{y})$ satisfies the x-transform condition on A with the exception of a progressively distributed set E;*

(c) *$f(x, y, \dot{y})$ satisfies condition* (2.4) *on E;*

(d) *K_a is a complete class of a.c. curves lying in A;*

then in the class K_a there is a minimizing curve for $I[y]$.

Since the projection of N has measure 0, there is a summable function $\phi(x)$, l.s.c. on $[a, b]$, such that $\phi(x) > 0$ on $[a, b]$ and

$$\phi(x) = \infty \tag{7.2}$$

on the projection of N. (De la Vallée Poussin's "majorante" for the function ∞ on the projection of N and 0 elsewhere will serve for $\phi(x)$.) Since $\nu + \phi$ is summable, it is easy to show that there is a positive-valued monotonic increasing continuous function $\mu(z)$ defined for $0 \leqq z$ such that

$$\lim_{z\to\infty} \mu(z)/z = \infty$$

and

$$\lambda(x) \equiv \mu(\nu(x) + \phi(x)) \tag{7.3}$$

is summable on $[a, b]$. Since μ is monotone and continuous, we see that $\lambda(x)$ is l.s.c. We define

$$\Lambda(x) = \int_a^x \lambda(x)\,dx,$$

and make the transformation $\xi = \Lambda(x)$. This is a.c. on $[a, b]$, and since $\lambda > 0$ it has an a.c. inverse $x = \Lambda^{-1}(\xi)$ $(0 \leqq \xi \leqq \Lambda(b))$. Under this transformation A goes over into the set A_1, consisting of those points (ξ, y) such that $(\Lambda^{-1}(\xi), y) \in A$; likewise E goes into a set E_1 which is progressively distributed, by Lemma 9, and N goes into a set N_1 whose projection on the ξ-axis has measure 0.

From (c) and (a) it follows that there is a number $c > 0$ such that $g_1(x, y, \dot{x}, \dot{y}) \equiv g(x, y, \dot{x}, \dot{y}) - c\dot{x}$ is monotonic decreasing[22] in $\dot{x}$. We now define a function $\bar{g}(\xi, y, \dot{\xi}, \dot{y})$ as follows: On the set H, consisting of all $(\xi, y, \dot{\xi}, \dot{y})$ with $(\xi, y) \in A_1 - N_1$ and $\dot{\xi} \geqq 0$ and all $(\xi, y, \dot{\xi}, \dot{y})$ with $(\xi, y) \in A_1$ and $\dot{y} = 0$, we set

$$\bar{g}(\xi, y, \dot{\xi}, \dot{y}) = g(x, y, \dot{\xi}/\lambda(x), \dot{y}) - c\dot{\xi}/\lambda(x), \tag{7.4}$$

wherein $x = \Lambda^{-1}(\xi)$. On the set K consisting of all $(\xi, y, \dot{\xi}, \dot{y})$ with $(\xi, y) \in A_1N_1$ and $|\dot{y}| \neq 0$ we set

$$\bar{g}(\xi, y, \dot{\xi}, \dot{y}) = \infty. \tag{7.5}$$

The function thus defined is l.s.c. Suppose $(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0) \in H + K$ and that $(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n) \in H + K$ and tends to $(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0)$. In calculating lim inf

[22] Cf. SC, p. 613.

$\bar{g}(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n)$ we may disregard all n such that $(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n) \epsilon K$, for $\bar{g} = \infty$ for all such arguments. So we assume $(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n) \epsilon H$. Define $x_i = \Lambda^{-1}(\xi_i)$ $(i = 0, 1, 2, \cdots)$. We consider two cases.

Case 1. $(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0) \epsilon K$. Then $(\xi_0, y_0) \epsilon N_1$, and $\nu(x_0) + \phi(x_0) = \infty$. Hence $\lim (\nu(x_n) + \phi(x_n)) = \infty$, and

$$(7.6) \qquad 0 \leqq \lim \nu(x_n)/\lambda(x_n) \leqq \lim (\nu(x_n) + \phi(x_n))/\lambda(x_n) = 0.$$

Therefore, $\dot{\xi}_n \nu(x_n)/\lambda(x_n) \to 0$, and by hypothesis (c)

$$\begin{aligned} \lim \bar{g}(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n) &= \lim \{g(x_n, y_n, [\dot{\xi}_n \nu(x_n)/\lambda(x_n)]/\nu(x_n), \dot{y}_n) - c\dot{\xi}_n/\lambda(x_n)\} \\ &= \infty - 0 = \bar{g}(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0). \end{aligned}$$

Case 2. $(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0) \epsilon H$. We first select a subsequence $\{p\}$ of the integers n such that

$$(7.7) \qquad \lim \bar{g}(\xi_p, y_p, \dot{\xi}_p, \dot{y}_p) = \liminf \bar{g}(\xi_n, y_n, \dot{\xi}_n, \dot{y}_n),$$

and then we select a subsequence $\{r\}$ of the p's such that

$$\lim \lambda(x_r) = \liminf \lambda(x_p).$$

Since $g - c\dot{x}$ is l.s.c. and is monotonic decreasing in $\dot{x}$, and $\lambda(x)$ is l.s.c.,

$$\begin{aligned} \lim \bar{g}(\xi_p, y_p, \dot{\xi}_p, \dot{y}_p) &= \lim \bar{g}(\xi_r, y_r, \dot{\xi}_r, \dot{y}_r) \\ &= \lim [g(x_r, y_r, \dot{\xi}_r/\lambda(x_r), \dot{y}_r) - c\dot{\xi}_r/\lambda(x_r)] \\ &\geqq g(x_0, y_0, \dot{\xi}_0/\lim \lambda(x_r), \dot{y}_0) - c\dot{\xi}_0/\lim \lambda(x_r) \\ &\geqq g(x_0, y_0, \dot{\xi}_0/\lambda(x_0), \dot{y}_0) - c\dot{\xi}_0/\lambda(x_0) \\ &= \bar{g}(\xi_0, y_0, \dot{\xi}_0, \dot{y}_0). \end{aligned}$$

This, with (7.7), establishes our conclusion.

If $(\xi_0, y_0) \epsilon A_1 - E_1$ and $|\dot{y}_0| \neq 0$, we find

$$(7.8) \qquad \bar{g}(\xi_0, y_0, 0, \dot{y}_0) = \infty.$$

For $(\xi_0, y_0, 0, \dot{y}_0) \epsilon K$ this is obvious from (7.5). For $(\xi_0, y_0, 0, \dot{y}_0) \epsilon H$, in the statement of the x-transform condition we take $x_n = x_0 = \Lambda^{-1}(\xi_0)$, $y_n = y_0$, $\dot{y}_n = \dot{y}_0$ for all n. Then $(x_n, y_n) \epsilon A - N$ by the definition of H, and by the x-transform condition

$$\bar{g}(\xi_0, y_0, 0, \dot{y}_0) = g(x_0, y_0, 0, \dot{y}_0) = \lim g(x_n, y_n, 0, \dot{y}_n) = \infty.$$

The function $\bar{g}(\xi, y, \dot{\xi}, \dot{y})$ satisfies condition (2.3c). For if $(\xi, y) \epsilon A_1 N_1$, then $g = \infty$ except along the $\dot{\xi}$-axis, where it is a linear function $d\dot{\xi}$. Then for any $(\dot{\xi}_0, \dot{y}_0)$ the linear function $a\dot{\xi} + b_\alpha \dot{y}^\alpha \equiv d\dot{\xi} + n\dot{y}_0^\alpha \dot{y}^\alpha$ will serve in (2.3c) if n is large enough. If $(\xi, y) \epsilon A_1 - N_1$ and $\lambda(\Lambda^{-1}(\xi)) \neq \infty$, then $\bar{g}(\xi, y, \dot{\xi}, \dot{y})$ is merely the linear transform of the convex function $g - c\dot{x}$, so it is convex and (2.3c) holds. If $(\xi, y) \epsilon A_1 - N_1$ and $\lambda(\Lambda^{-1}(\xi)) = \infty$, then $\bar{g}(\xi, y, \dot{\xi}, \dot{y}) = g(x, y, 0, \dot{y})$, which is convex in $(\dot{\xi}, \dot{y})$.

Moreover, if $(\xi, y) \in E_1$, condition (2.6) holds. For then there is a linear function $a\dot{x} + b_\alpha \dot{y}^\alpha$ such that

$$g(\Lambda^{-1}(\xi), y, \dot{x}, \dot{y}) > a\dot{x} + b_\alpha \dot{y}^\alpha \quad \text{for } (\dot{x}, \dot{y}) \neq (0, 0),$$

by hypothesis (c) and Lemma 1. Then the linear function $[-1 + (a - c)/\lambda(x)]\dot{\xi} + b_\alpha \dot{y}^\alpha$ serves for $\bar{g}$. For either $(\xi, y, \dot{\xi}, \dot{y}) \in H$, and then

$$\begin{aligned} \bar{g}(\xi, y, \dot{\xi}, \dot{y}) &= g(x, y, \dot{\xi}/\lambda(x), \dot{y}) - c\dot{\xi}/\lambda(x) \\ &> (a - c)\dot{\xi}/\lambda(x) + b_\alpha \dot{y}^\alpha \quad \text{for } (\dot{\xi}/\lambda(x), \dot{y}) \neq (0, 0), \end{aligned}$$

whence

$$\bar{g}(\xi, y, \dot{\xi}, \dot{y}) > [-1 + (a - c)/\lambda(x)]\dot{\xi} + b_\alpha \dot{y}^\alpha \quad \text{for } (\dot{\xi}, \dot{y}) \neq (0, 0);$$

or else $(\xi, y, \dot{\xi}, \dot{y}) \in K$, and

$$\bar{g}(\xi, y, \dot{\xi}, \dot{y}) = \infty > [-1 + (a - c)/\lambda(x)]\dot{\xi} + b_\alpha \dot{y}^\alpha.$$

Now let $C: y = y(x)$ $(x_1 \leqq x \leqq x_2)$ be an a.c. curve in A, and let

$$\Gamma: y = \eta(\xi) \equiv y(\Lambda^{-1}(\xi)), \qquad \xi_1 \equiv \Lambda(x_1) \leqq \xi \leqq \Lambda(x_2) \equiv \xi_2$$

be its transform. Then

$$\begin{aligned} \bar{I}[\eta] &= \int_{\xi_1}^{\xi_2} \bar{g}(\xi, \eta, 1, \dot{\eta})\, d\xi \\ &= \int_{x_1}^{x_2} \bar{g}(\Lambda(x), \eta(\Lambda(x)), 1, \dot{\eta}(\Lambda(x)))\cdot\dot{\Lambda}(x)\, dx \\ &= \int_{x_1}^{x_2} \bar{g}(\Lambda(x), y(x), 1, \dot{y}(x)/\lambda(x))\lambda(x)\, dx \\ &= \int_{x_1}^{x_2} \bar{g}(\Lambda(x), y(x), \lambda(x), \dot{y}(x))\, dx \\ &= \int_{x_1}^{x_2} [g(x, y(x), 1, \dot{y}(x)) - c]\, dx \\ &= I[y] - c(x_2 - x_1). \end{aligned} \tag{7.9}$$

If $C_n: y = y_n(x)$ $(a_n \leqq x \leqq b_n)$ is a sequence of a.c. curves in A such that the numbers $I[y_n]$ are bounded, then by (7.9) for the transforms $\Gamma_n: y = \eta_n(\xi)$ the numbers $\bar{I}[\eta_n]$ are bounded. By Lemmas 13 and 3, the curves Γ_n have uniformly bounded lengths. They therefore have a curve of accumulation Γ_0. By our semi-continuity theorem[23] $\bar{J}[\Gamma_0] \leqq \liminf \bar{I}[\eta_n] < \infty$, so by Lemma 12 Γ_0 has an a.c. representation $y = \eta_0(\xi)$. The transformation from A to A_1 is continuous both ways, so the curves C_n have for their curve of accumulation the transform C_0 of Γ_0. This transform is given by the equation $C_0: y = y_0(x) \equiv$

[23] SC, Theorem 3.1. The function $\bar{g}$ satisfies conditions 2.2 of SC.

$\eta_0(\Lambda(x))$, and since the last function is an a.c. function of a monotonic a.c. function, it is a.c.

In particular, if the curves C_n form a minimizing sequence in the class K_a, then the curve C_0 is in K_a, because K_a is complete. So by the usual argument C_0 is a minimizing curve for $I[y]$ on the class K_a.

8. The x-transform condition defined in the preceding section is defined with respect to the whole set A. We shall now study local criteria which will ensure that the x-transform condition is satisfied on A.

LEMMA 14. *If A is bounded and closed, and if at all points (x_0, y_0) of A except at most those of a progressively distributed set E_0 one of the two conditions holds:*

(i) $g(x_0, y_0, 0, \dot{y}) = \infty$ *for* $|\dot{y}| \neq 0$;

(ii) *there is a neighborhood* $U: |x - x_0| < \epsilon, |y - y_0| < \epsilon$ *such that g satisfies the x-transform condition on AU with the exception of a progressively distributed set E;*

then g satisfies the x-transform condition on A with the exception of a progressively distributed set.

Let B be the subset of A on which (i) holds, and let $C = A - B - E_0$. To each (x, y) in C there is a neighborhood U as described in (ii). A denumerable collection of these neighborhoods covers C. Denote these neighborhoods by $U_i: |x - \bar{x}_i| < \epsilon_i, |y - \bar{y}_i| < \epsilon_i, (i = 1, 2, \cdots)$. For each i there is a set N_i whose projection on the x-axis has measure 0, a progressively distributed set E_i and a positive l.s.c. function $\nu_i(x)$ summable over $[\bar{x}_i - \epsilon_i, \bar{x}_i + \epsilon_i]$ which serves in the definition of the x-transform condition on AU. Define

$$E = \sum_0^\infty E_i, \qquad N = \sum_1^\infty N_i,$$

$$k_i = 2^{-i} \Big/ \int_{\bar{x}_i - \epsilon_i}^{\bar{x}_i + \epsilon_i} \nu_i(x)\, dx,$$

$$\nu_i^*(x) = \begin{cases} 0 & \text{if} \quad x \leqq \bar{x}_i - \epsilon_i \quad \text{or} \quad x \geqq \bar{x}_i + \epsilon_i, \\ k_i \nu_i(x) & \text{if} \quad \bar{x}_i - \epsilon_i < x < \bar{x}_i + \epsilon_i, \end{cases}$$

$$\nu(x) = 1 + \sum_1^\infty \nu_i(x).$$

Let a, b be numbers such that A lies between $x = a$ and $x = b$. Then

$$\int_a^b \nu_i^*(x)\, dx = 2^{-i},$$

so that $\nu(x)$ is summable on (a, b). Also $\nu(x)$ is l.s.c., being the sum of non-negative l.s.c. functions. The set E is progressively distributed by Lemma 6, and the projection of N on the x-axis has measure 0.

Now let (x_0, y_0) be in $A - E$, and $|\dot{y}_0| \neq 0$, and let $(x_n, y_n, \dot{x}_n, \dot{y}_n)$ be a

sequence tending to $(x_0, y_0, 0, \dot{y}_0)$ and such that $(x_n, y_n) \in A - N$. We must prove

$$\lim_{n\to\infty} g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) = \infty. \tag{8.1}$$

Suppose first that $(x_0, y_0) \in C$. Then $(x_0, y_0) \in U_i$ for some i. If we set

$$v_n = \dot{x}_n \nu_i(x_n)/\nu(x_n),$$

then since $\nu(x) > k_i\nu_i(x)$ we see that

$$0 \leqq \lim_{n\to\infty} v_n \leqq \lim_{n\to\infty} \dot{x}_n/k_i = 0.$$

For large n the point $(x_n, y_n) \in U_i$, and we have assumed

$$(x_n, y_n) \in A - N \subset A - N_i \quad \text{and} \quad (x_0, y_0) \in AU_i - E \subset AU_i - E_i.$$

So by the definition of the x-transform condition

$$\lim_{n\to\infty} g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) = \lim_{n\to\infty} g(x_n, y_n, v_n/\nu_i(x_n), \dot{y}_n) = \infty.$$

Suppose next that $(x_0, y_0) \in B$. Then by the semi-continuity of g, and the relation $\dot{x}_n/\nu(x_n) \to 0$ we have

$$\liminf_{n\to\infty} g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) \geqq g(x_0, y_0, 0, \dot{y}_0) = \infty.$$

Thus (8.1) holds in all cases, and the lemma is established.

LEMMA[24] 15. *The function g satisfies the x-transform condition on a neighborhood of the point (x_0, y_0) of A if it satisfies the following condition (condition β). There exist three functions $\phi(z)$, $\psi(z)$, $M(z)$ and three constants $l > 0$, $a > 0$, μ such that*

(i) *$\phi(z)$ is defined and summable in $(0, l)$, and $\phi(z) \to \infty$ as $z \to 0$;*

(ii) *$\psi(z)$ is defined and continuous*[25] *on $(0, \infty)$, is non-negative and monotonic increasing, and*

$$\liminf_{z\to 0} z\phi(z)\psi(\phi(z)) > 0;$$

(iii) *$M(z)$ is defined for $z \geqq 0$, and $M(z) \to \infty$ as $z \to \infty$;*

(iv) *the inequality*

$$f(x, y, \dot{y}) \geqq |\dot{y}|\, M(|x - x_0|\cdot|\dot{y}|\cdot\psi(|\dot{y}|)) + \mu \tag{8.2}$$

holds for all $(x, y) \in A$ on a neighborhood U of (x_0, y_0).

To begin with, we may assume that $\phi(z)$ is l.s.c., for if we define $\gamma(z)$ to be the lower limit function of $\phi(z)$ (that is, the smaller of $\phi(z)$ and $\liminf \phi(\bar{z})$ as $\bar{z} \to z$), then $\gamma(z)$ is l.s.c. There is a $k > 0$ such that $z\phi(z)\,\psi(\phi(z)) \geqq k$ if z is small, by (iii). For such z, choose a sequence $z_n \to z$ such that $\phi(z_n) \to \gamma(z)$; then

$$z\gamma(z)\,\psi(\gamma(z)) = \lim z_n\phi(z_n)\,\psi(\phi(z_n)) \geqq k.$$

[24] Tonelli, op. cit. (footnote 2), p. 17.

[25] The assumption of continuity is not really needed.

So γ would serve in place of ϕ, and we therefore may suppose ϕ to be l.s.c. from the start. Likewise, since $\lim_{z \to 0} \phi(z) = \infty$, by restricting z to a neighborhood $[0, l']$ with $l' < l$ we can ensure $\phi(z) > 1$. Also, by shrinking U, we can obtain $| x - x_0 | < l'$ for all $(x, y) \in U$.

Now in the definition of the x-transform condition take E empty and N consisting of the plane $x = x_0$, and let $\nu(x) = \phi(| x - x_0 |)$. Consider any sequence $(x_n, y_n, \dot{x}_n, \dot{y}_n)$ with $(x_n, y_n) \in AU - N$ (that is, $x_n \neq x_0$) and $\dot{x}_n \to 0$, $\dot{x}_n > 0$. By our usual definition (1.2) of g and inequality (8.2)

$$\begin{aligned}(8.3) \qquad g(x_n, y_n, \dot{x}_n, \dot{y}_n) &= \dot{x}_n f(x_n, y_n, \dot{y}_n/\dot{x}_n) \\ &\geqq | \dot{y}_n | M(| x_n - x_0 | \cdot (| \dot{y}_n |/\dot{x}_n) \cdot \psi(| \dot{y}_n |/\dot{x}_n)) + \mu | \dot{x}_n | .\end{aligned}$$

Let $(x, y) \in AU - N$, $| \dot{y} | \neq 0$, and take $x_n \equiv x$, $y_n \equiv y$, $\dot{x}_n = 1/n$, $\dot{y}_n = \dot{y}$ for all n. Then by definition

$$\begin{aligned}g(x, y, 0, \dot{y}) &= \lim g_n(x_n, y_n, \dot{x}_n, \dot{y}_n) \\ &\geqq \lim | \dot{y} | \cdot M(| x - x_0 | \cdot n \cdot | \dot{y} | \cdot \psi(n | \dot{y} |)).\end{aligned}$$

By (ii), $\psi(n | \dot{y} |)$ has a positive lower bound for large n, so by (iii) the limit on the right is ∞, and so

$$(8.4) \qquad g(x, y, 0, \dot{y}) = \infty \quad \text{if } (x, y) \in AU - N \text{ and } | \dot{y} | \neq 0.$$

Now let $(\bar{x}, y_0) \in AU$, $| \dot{y}_0 | \neq 0$ and let $(x_n, y_n, \dot{x}_n, \dot{y}_n)$ be a sequence tending to $(\bar{x}, y_0, 0, \dot{y}_0)$ and having $(x_n, y_n) \in AU - N$ (i.e., $x_n \neq x_0$). We wish to show

$$(8.5) \qquad \lim g(x_n, y_n, \dot{x}_n/\phi(| x_n - x_0 |), \dot{y}_n) = \infty .$$

If $\bar{x} \neq x_0$, this follows at once from (8.4) and the lower semi-continuity of g. If $\bar{x} = x_0$, we observe that in the subsequence (if any) for which $\dot{x}_n/\phi(x_n) = 0$ the relation (8.5) is obvious by (8.4). So we suppose $\dot{x}_n/\phi(x_n) > 0$. Also we disregard the terms (finite in number) for which $\dot{x}_n > | \dot{y}_n |$. Then by (8.3) and (ii)

$$\begin{aligned}&g(x_n, y_n, \dot{x}_n/\phi(| x_n - x_0 |), \dot{y}_n) \\ &\quad \geqq | \dot{y}_n | M\left(| x_n - x_0 | \frac{| \dot{y}_n | \phi(| x_n - x_0 |)}{\dot{x}_n} \psi\left(\frac{| \dot{y}_n | \phi(| x_n - x_0 |)}{\dot{x}_n}\right)\right) + \frac{\mu \dot{x}_n}{\phi(| x_n - x_0 |)} \\ &\quad \geqq | \dot{y}_n | M\left(\frac{\dot{y}_n}{\dot{x}_n} | x_n - x_0 | \phi(| x_n - x_0 |) \psi(\phi(| x_n - x_0 |))\right) + \frac{\mu \dot{x}_n}{\phi(| x_n - x_0 |)}.\end{aligned}$$

In the argument of M the factor $| \dot{y}_n |/\dot{x}_n$ tends to ∞, and the other factor has a positive lower limit by (ii), so by (iii) the first term on the right tends to ∞. The second term tends to 0. This establishes (8.5). This completes the proof that g satisfies the x-transform condition on AU.

COROLLARY. *Condition (β) is satisfied if ϕ is positive, continuous and monotonic*

decreasing in the strict sense on $(0, l)$, *and* $z\phi^{-1}(z)$ *is monotonic decreasing on* $(0, \infty)$, *and*

$$(8.6) \qquad f(x, y, \dot{y}) \geqq |\dot{y}| M(|x - x_0|/\phi^{-1}(|\dot{y}|)) + \mu$$

for $(x, y) \epsilon AU$, *where* M *satisfies* (iii) *of Lemma* 15.

Define $\psi(z) = 1/z\phi^{-1}(z)$; then (ii) holds, and (iv) takes the form (8.6).

This corollary enables us to set up a whole array of fairly simple tests; for we can take $M(z)$ to be $kz^\alpha (\alpha > 0)$, or $k \log(1 + z)$, or $k \log(\log(z + e))$ $(k > 0)$, etc., and then choose a function $\sigma(t) > 0$, continuous and monotonic increasing on the interval $0 < t$, and such that $1/\sigma(t)$ is summable from 0 to ∞. Then if

$$f(x, y, \dot{y}) \geqq |\dot{y}| M(|x - x_0| \sigma(|\dot{y}|)) \quad \text{for}(x, y) \epsilon AU,$$

the hypotheses of the corollary hold with $\phi^{-1}(z) = 1/\sigma(z)$.

9. **Extension to unbounded fields.** In this section we shall use a function $\phi(x, y^1, \cdots, y^q)$ of class C'. For ease in printing, its partial derivatives with respect to $x, y^1, \cdots, y^q$ shall be respectively denoted by $\phi_0, \phi_1, \cdots, \phi_q$.

We shall now prove[26]

LEMMA 16. *If*

(a) A *is a closed set lying between two hyperplanes* $x = -c$ *and* $x = c$, *and* A_0 *is a bounded closed subset of* A;

(b) $\phi(x, y)$ *is of class* C' *on* A, *and tends to* ∞ *as* $|y| \to \infty$, *uniformly in* x;

(c) K_a *is a class of absolutely continuous curves* $y = y(x)$ *in* A, *each having at least one point in common with* A_0;

(d) *for all* (x, y) *in* $A - A_0$ *and all* y' *the function* $f(x, y, \dot{y})$ *satisfies the condition*

$$(9.1) \qquad f(x, y, \dot{y}) \geqq |\phi_0(x, y) + \phi_\alpha(x, y)\dot{y}^\alpha|;$$

then for every M *the class of curves of* K_a *with* $I[y] \leqq M$ *lies in a bounded subset of* A.

Let N be a lower bound for $f(x, y, \dot{y})$. Then if we replace f by $f - N$ and $\phi(x, y)$ by $\phi(x, y) - Nx$, all hypotheses remain satisfied, and the new f is non-negative. So we may suppose $f \geqq 0$.

Let m be an upper bound for $\phi(x, y)$ on A_0. By (b), there exists an r such that

$$(9.2) \qquad \phi(x, y) > M + m \qquad \text{whenever } (x, y) \epsilon A \text{ and } |y| > r.$$

We are at liberty to assume that $|y| < r$ whenever $(x, y) \epsilon A_0$, since we can increase r if we will.

Now consider a curve $y = y(x)$ $(a \leqq x \leqq b)$ of the class K_a. Let $(x_0, y(x_0))$ be a point of the curve. If $(x_0, y(x_0)) \epsilon A_0$, then $|y_0| < r$. If not, either

[26] The lemma and its proof are essentially due to Cinquini, loc. cit. (footnote 4).

there is a greatest $x_1 < x_0$ such that $(x_1, y(x_1)) \epsilon A_0$, or else there is a least $x_1 > x_0$ for which this holds. Suppose the former, for definiteness. Then

$$(9.3)\quad \begin{aligned} M &\geqq \int_a^b f(x, y, \dot{y})\,dx \geqq \int_{x_1}^{x_0} f(x, y, \dot{y})\,dx \geqq \int_{x_1}^{x_0} \{\phi_0(x, y) + \phi_\alpha(x, y)\dot{y}^\alpha\}\,dx \\ &= \phi(x_0, y(x_0)) - \phi(x_1, y(x_1)), \end{aligned}$$

so $\phi(x_0, y(x_0)) \leqq M + m$ and $| y(x_0) | \leqq r$. Thus in either case $(x_0, y(x_0))$ lies in the bounded set common to A and the cylinder $| y | \leqq r$. This establishes the lemma.

As a corollary, if the initial point of each curve of K lies in A_0, we can replace (9.1) by

$$(9.4)\quad f(x, y, \dot{y}) \geqq \phi_0(x, y) + \phi_\alpha(x, y)\dot{y}^\alpha.$$

For then in (9.3) there surely is an $x_1 \leqq x_0$ such that $(x_1, y(x_1)) \epsilon A_0$. Likewise, if the final point of each curve of K is in A, then we can replace (9.1) by

$$(9.5)\quad f(x, y, \dot{y}) \geqq -\phi_0(x, y) - \phi_\alpha(x, y)\dot{y}^\alpha.$$

Also, we could require only that there exist a constant h such that (9.1) holds whenever $| \phi_0 + \phi_\alpha \dot{y}^\alpha | \geqq h$; for then (9.1) holds with f replaced by $f + h$, and the class of curves of K with $\int f dx \leqq M$ is contained in the class with $\int (f + h)\,dx \leqq M + 2hc$, which by Lemma 16 lies in a bounded subset of A.

As an example, we can take $\phi = k \log (1 + y^\alpha y^\alpha)$ $(k > 0)$. Condition (9.1) then becomes

$$f(x, y, \dot{y}) \geqq 2k(1 + y^\alpha y^\alpha)^{-1} (y^\alpha \dot{y}^\alpha),$$

which is surely satisfied if

$$f(x, y, \dot{y}) \geqq 2k \,| \dot{y} |/| y |\,.$$

We need only ask that this hold when $| y |$ exceeds a constant h, for we can add to A_0 the part of A in the cylinder $| y | \leqq h$.

10. From the preceding existence theorems we readily deduce an existence theorem for solutions of problems involving higher derivatives,

$$(10.1)\quad I \equiv \int f(x, u, u', \cdots, u^{(h)}, z, z', \cdots, z^{(k)}, \cdots, w, w', \cdots, w^{(l)})\,dx = \min.$$

Let us change notation. We denote the arguments $u, u', \cdots, u^{(h-1)}, z, z', \cdots, z^{(k-1)}, \cdots, w, w', \cdots, w^{(l-1)}$ by $y^1, y^2, \cdots, y^p$, respectively. For certain values of i the symbol y^i denotes an end of a sequence such as $u, u', \cdots, u^{(h-1)}$ or $w, w', \cdots, w^{(l-1)}$. This collection of values of i we denote by Δ; the rest of the numbers $1, \cdots, p$ form the complementary class Δ'. To save repetition, the symbol i shall be used for integers in Δ and the symbol j for integers in Δ'. The

space determined by x and the y^i $(i \in \Delta)$ will be denoted by "(x, y_Δ)-space". Our problem is then that of minimizing in a certain class K of curves, the value of

$$(10.2) \qquad I[y] \equiv \int_a^b f(x, y, \dot{y})\, dx,$$

subject to the conditions:
(10.3) f is independent of the $\dot{y}^j$;

$$(10.4) \qquad y^j(x) = y^j(a) + \int_a^x y^{j+1}(x)\, dx \qquad (a \leqq x \leqq b,\ j \in \Delta').$$

For such problems we shall establish

THEOREM 5. *If*

(a) *A is a bounded closed point-set in (x, y)-space;*

(b) *$f(x, y, \dot{y})$ satisfies conditions* (2.1) *and* (10.3) *on A;*

(c) *there is a subset E of A, whose projection on (x, y_Δ)-space is progressively distributed, such that*

(i) *if $(x, y) \in E$, there is a linear function $a\dot{x} + \sum_{i \in \Delta} b_i \dot{y}^i$ such that*

$$g(x, y, \dot{x}, \dot{y}) + a\dot{x} + \sum b_i \dot{y}^i > 0 \quad \text{if } (\dot{x}, \dot{y}^i) \neq (0, \cdots, 0);$$

(ii) *g satisfies the x-transform condition (with the condition $|\dot{y}| \neq 0$ replaced by $(\dot{y}^i) \neq (0, \cdots, 0)$) on A with the exception of E;*

(d) *$f(x, y, 0)$ is bounded on A;*

(e) *K_a is a complete class of a.c. curves lying in A;*

then among those curves of K_a which satisfy (10.4) *there exists a minimizing curve for $I[y]$.*

To begin with we prove that the class D of curves satisfying (10.4) is complete. Suppose that the a.c. curves

$$C_k: y = y_k(x) \qquad (a_k \leqq x \leqq b_k;\ k = 0, 1, 2, \cdots)$$

are such that $C_n \in D$ $(n = 1, 2, \cdots)$ and $C_n \to C_0$; we must show $C_0 \in D$. If $a_0 = b_0$ this is trivial. If not, let α, β be numbers such that $a_0 < \alpha < \beta < b_0$. Since $a_n \to a_0$ and $b_n \to b_0$, the interval (α, β) is interior to (a_n, b_n) for almost all n. From (10.4) we deduce

$$(10.5) \qquad y_n^j(x) = y_n^j(\alpha) + \int_\alpha^x y_n^{j+1}(x)\, dx \qquad (\alpha \leqq x \leqq \beta,\ j \in \Delta').$$

Since $C_n \to C_0$, the functions y_n^{j+1} converge uniformly to y_0^{j+1}, so that

$$(10.6) \qquad y_0^j(x) = y_0^j(\alpha) + \int_\alpha^x y_0^{j+1}(x)\, dx \qquad (\alpha \leqq x \leqq \beta,\ j \in \Delta').$$

The functions y_0 are continuous, so we may let $\alpha \to a_0$ and $\beta \to b_0$ and obtain equation (10.4) for y_0.

The class K_aD being complete, we could apply Theorem 4 and obtain our conclusion except for two difficulties. One is that g does not satisfy the x-

transform condition except in our altered form; there is a summable l.s.c. function $\nu(x) > 0$ such that

$$g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) \to \infty$$

if $(x_n, y_n) \in A - N \to (x_0, y_0) \in A - E$ and $\dot{x}_n \to 0$ and $\dot{y}_n \to \dot{y}_0$, where the $\dot{y}_0^i$ are not all 0. The second difficulty is that E is not necessarily progressively distributed.

The first difficulty is easily remedied. If we define

$$F(x, y, \dot{y}) = f(x, y, \dot{y}) + \sum_{j \in \Delta'} (\dot{y}^j - y^{j+1})^2,$$

then for all curves satisfying (10.4) the integrals of F and of f are the same, since the added terms vanish identically along such curves. For the associated integrand $G(x, y, \dot{x}, \dot{y})$ we find

$$\begin{aligned}(10.7)\quad G(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n)& \\ = g(x_n, y_n, \dot{x}_n/\nu(x_n), \dot{y}_n) &+ \sum_{j \in \Delta'} (\dot{y}_n^j \nu(x_n)/\dot{x}_n - y_n^{j+1})^2.\end{aligned}$$

Suppose now that $(x_n, y_n) \in A - N$ and $(x_n, y_n, \dot{x}_n, \dot{y}_n) \to (x_0, y_0, 0, \dot{y}_0)$, where $(x_0, y_0) \in A - E$ and $|\dot{y}_0| \neq 0$. If the $\dot{y}_0^i$ are not all zero, the first term on the right tends to ∞ by hypothesis, while the other terms are non-negative. If the $\dot{y}_0^i$ are all zero, some $\dot{y}_0^j \neq 0$. Thus in the corresponding term on the right we have $\dot{y}_n^j \to \dot{y}_0^j \neq 0$, $\liminf \nu(x_n) \geqq \nu(x_0) > 0$, $\dot{x}_n \to 0$; so that term tends to ∞, while the other terms are bounded below. In either case, the left member of (10.7) tends to ∞, and G satisfies the x-transform condition on A with the exception of E.

Now let $y = y_n(x)$ $(a_n \leqq x \leqq b_n)$ be a minimizing sequence for $I[y]$. Since $|y_n(x)|$ is uniformly bounded, by (10.4) the $y_n^j(x)$ satisfy a Lipschitz condition uniformly, and we can therefore select a subsequence (which we may suppose to be the whole sequence) such that the $y_n^j(x)$ converge to limit functions $y_0^j(x)$ $(a_0 \leqq x \leqq b_0)$. Let Γ be the set of all points (x, y) such that $a_p \leqq x \leqq b_p$ and $y^j = y_p^j(x)$, $j \in \Delta'$, for some p of the set $0, 1, 2, \cdots$. Then Γ is easily seen to be closed; and its projection on each of the planes (x, y^j) is progressively distributed, being composed of a denumerable set of a.c. curves. The lower bound of $I[y]$ on the class K_aD is the same as its lower bound on the set K_aD_Γ of those curves of K_aD which lie in $A\Gamma$, because the minimizing sequence $y = y_n(x)$ lies in $A\Gamma$. The set $E\Gamma$ has progressively distributed projections on the planes (x, y^i) and on the (x, y_Δ)-space, so by Lemma 7 $E\Gamma$ is progressively distributed.

Now the point-set $A\Gamma$, the class K_aD_Γ and the integral $\int F dx$ satisfy the hypotheses of Theorem 4, so there is an a.c. curve $y = y(x)$ $(a \leqq x \leqq b)$ for which $\int F dx$ attains its lower bound in the class K_aD_Γ. But then for $y = y(x)$ the integral $\int F dx$ also attains its lower bound in the class K_aD. On this class $\int F dx$ is identical with $I[y]$, so $I[y]$ assumes its lower bound on the class K_aD, as was to be proved.

If in f only first derivatives enter, Theorem 5 reduces exactly to Theorem 4.

The extension to unbounded fields A can be made without difficulty by use of §9. In this connection there is one especially interesting case. If we return to the notation of (10.1), the class K_a may consist of all a.c. curves $u = u(x), \cdots, w = w(x)$ joining two fixed points $(x_1, u_1, \cdots, w_1)$ and $(x_2, u_2, \cdots, w_2)$, the derivatives at those end-points being unrestricted. This is in fact a problem with variable end-points; and even if $(x, u, \cdots, w)$ are restricted to lie in a bounded closed set B the range of (x, y) (in the notation of (10.2)) is unbounded,[27] for the values of $u', \cdots, u^{(h-1)}$, etc., are unbounded. However, if we impose the condition of §9 and also a condition[28] which will ensure that at least one set of arguments

$$(x, u, \cdots, u^{(h-1)}, \cdots, w, w', \cdots, w^{(l-1)})$$

of each admissible curve shall lie in a bounded set A_0, then the conclusion of Theorem 5 holds.

UNIVERSITY OF VIRGINIA.

[27] Excepting the case $h = k = \cdots = l = 1$.

[28] Such a condition, for example, is the following: $f(x, u, \cdots, u^{(h)}, \cdots, w, \cdots, w^{(l)}) \to \infty$ uniformly in $(x, u^{(h)}, \cdots, w^{(l)})$ as $\Sigma [u^2 + \cdots + (u^{(h-1)})^2 + \cdots + w^2 + \cdots + (w^{(l-1)})^2] \to \infty$.

TRANSFORMATIONS IN LINEAR TOPOLOGICAL SPACES

By John V. Wehausen

1. **Introduction.** In this paper some well known theorems for linear metric spaces whose proofs do not involve completeness are extended to linear topological spaces. The relation between bounded sets and linear continuous transformations is considered; and, in connection with this, metrizability conditions for linear topological spaces are obtained in terms of boundedness. A linear continuous functional on a linear topological space is shown to operate essentially on only a "part" of the space which can be normed. Finally the relation between completeness and category is considered and a theorem is proved which shows that two definitions of completeness which have been given do not imply that the space is of the second category.

2. **Topologies for linear spaces.** In the postulates for a linear set [cf. **1**, p. 26][1] there is no notion of a topology. If a topology is imposed on a linear set in such a fashion that the postulated operations of addition of elements and multiplication of elements by real numbers are continuous in the topology, the linear set will be said to be a linear topological space [cf. **2**, pp. 201–204]. The most important topologies of this type for linear sets have been metric topologies, viz., the F-metric [**1**, p. 35] and the norm [**1**, p. 53].[2] As a matter of fact, it will be shown that the F-metric is the most general metric of this nature.

However, two non-metric topologies have been given in the literature. Kolmogoroff [**3**, p. 29] has simply postulated an operation of closure, $\bar{M}$, for any set M satisfying the Riesz-Kuratowski axioms, viz.,

(1) if M is a single element, $\bar{M} = M$,

(2) $\bar{\bar{M}} = \bar{M}$,

(3) $\overline{M + N} = \bar{M} + \bar{N}$,

and then in addition required that $x + y$ and αx be continuous in the topology. He proves that such a space is a regular Hausdorff space. On the other hand, John von Neumann [**4**, p. 4] has defined a topology for a linear set L by means of a set of neighborhoods $\mathfrak{U} = \{U\}$ satisfying the following axioms:

(1) $\theta \in U$ for every $U \in \mathfrak{U}$.

(2) If $U, V \in \mathfrak{U}$, there exists $W \in \mathfrak{U}$ such that $W \subset \mathfrak{P}(U, V)$.

(3) If $U \in \mathfrak{U}$, there exists $V \in \mathfrak{U}$ such that $V + V \subset U$.

(4) If $U \in \mathfrak{U}$, there exists $V \in \mathfrak{U}$ such that $\alpha V \subset U$ for $|\alpha| \leqq 1$.

Received November 9, 1937.

[1] Boldface numbers refer to the bibliography at the end of the paper.

[2] Other non-continuous metrics have been introduced by C. R. Adams [**12**, pp. 422, 423].

(5) If $x \in L$, $U \in \mathfrak{U}$, there exists α such that $x \in \alpha U$.
L is called a convex space if $\mathfrak{U}$ satisfies in addition

(6) $U + U \subset 2U$ for all $U \in \mathfrak{U}$.

The neighborhood system for any element is then obtained by translating the neighborhoods for θ. It is proved by von Neumann that Axioms (1)-(5) give a regular Hausdorff space and that the operations of addition and multiplication are continuous [**4**, p. 6, Theorems 6 and 7]. This topology then satisfies Kolmogoroff's axioms. The converse is also true. As the neighborhood system for each element x of the space take the set of all open sets containing the element, denoting it by $\mathfrak{U}(x)$. First it must be shown that the topology is "uniform" over the whole space, i.e., if $x \in L$, then the translated neighborhood system $\mathfrak{U}(\theta) + x$ is equivalent to $\mathfrak{U}(x)$. Let $U_x \in \mathfrak{U}(x)$ be arbitrary. Since $\theta + x = x$, it follows from continuity of addition that there exist $V_\theta \in \mathfrak{U}(\theta)$ and $V_x \in \mathfrak{U}(x)$ such that $V_\theta + V_x \subset U_x$. Since $x \in V_x$, $V_\theta + x \subset U_x$. The definition of continuity in connection with the equation $x + (-x) = \theta$ gives the other half of the equivalence. Therefore the topology of Kolmogoroff is uniform. It then only remains to show that Axioms (1)-(5) are satisfied by $\mathfrak{U}(\theta)$. But (1) is satisfied by definition of $\mathfrak{U}(\theta)$ and (2) is a property of any Hausdorff space. Axiom (3) follows from continuity of addition as applied to $\theta + \theta = \theta$ and (4) from continuity of multiplication as applied to $0 \cdot \theta = \theta$. Consider now (5). Let $x \in L$ and $U \in \mathfrak{U}(\theta)$ be arbitrary and suppose that x non-$\in nU$, i.e., $n^{-1}x$ non-$\in U$ for every integer n. But since multiplication is continuous $n^{-1}x \to \theta$, i.e., there exists $n(U)$ such that, for $n \geqq n(U)$, $n^{-1}x \in U$. This is a contradiction. The following theorem has thus been established.

THEOREM 1. *The topologies of Kolmogoroff and von Neumann are equivalent.*

Tychonoff [**6**, p. 768] has called a linear topological space satisfying Kolmogoroff's axioms "locally convex" if for every U_x there exists a convex[3] neighborhood V_x such that $V_x \subset U_x$. It follows easily from the regularity of the space and a theorem of von Neumann [**4**, p. 10, Theorem 13] that the notions of "convex" and "locally convex" space are the same. One may therefore state the following corollary of the previous theorem.

COROLLARY. *The convex topologies of von Neumann are equivalent to the locally convex topologies of Tychonoff.*

3. **Preliminary definitions**. If L_1 and L_2 are two linear topological spaces, the definitions of additive and homogeneous transformations on L_1 to L_2 are unchanged. Continuity of transformations is defined as usual for neighborhood spaces. It is easily seen that an additive transformation which is continuous at one point is everywhere continuous and is homogeneous, so that one need define continuity for such transformations only at the zero-element, θ, i.e., if $\mathfrak{U} = \{U\}$ and $\mathfrak{V} = \{V\}$ are the neighborhood systems in L_1 and L_2 respectively, then $T(x)$ on L_1 to L_2 is continuous if and only if for arbitrary $V \in \mathfrak{V}$ there exists $U \in \mathfrak{U}$ such that $T(U) \subset V$.

[3] A set M is convex if, whenever $x, y \in M$, $\alpha, \beta \geqq 0$ and $\alpha + \beta = 1$, then $\alpha x + \beta y \in M$.

Mazur and Orlicz [7, p. 152] have used the following definition for a bounded set:[4] $R \subset L$ is bounded if for every sequence $\{x_i\} \subset R$ and every sequence of real numbers $\alpha_i \to 0$, the sequence $\alpha_i x_i \to \theta$. On the other hand, von Neumann [4, p. 7, Definition 5] uses this definition: $R \subset L$ is bounded if for every $U \in \mathfrak{U}$ there exists a real number α such that $R \subset \alpha U$. The following lemma is stated without proof.

LEMMA. *The definitions of boundedness of Banach and of von Neumann are equivalent.*[5]

4. **Relation between bounded sets and continuity of additive transformations.** In normed linear spaces the equivalence of the class of additive transformations taking bounded sets into bounded sets and the class of additive continuous transformations is a well known theorem. This theorem has been extended by Mazur and Orlicz [7, p. 153] to F-spaces. Half of the equivalence may be easily proved for linear topological spaces also.

THEOREM 2. *If $T(x)$ is a linear continuous transformation on L_1 to L_2 and if $R \subset L_1$ is bounded, then $T(R) \subset L_2$ is bounded.*

For let $V \in \mathfrak{V}$ be arbitrary. Then by continuity of $T(x)$ there exists $U \in \mathfrak{U}$ such that $T(U) \subset V$. Since R is bounded, there exists α such that $R \subset \alpha U$. Hence, $T(R) \subset T(\alpha U) = \alpha T(U) \subset \alpha V$, i.e., $T(R)$ is bounded since V was arbitrary.

The converse of this theorem may be proved in the following restricted form.

THEOREM 3. *If L contains a bounded open set and if the additive transformation $T(x)$ takes bounded sets into bounded sets, $T(x)$ is continuous.*

The proof will not be given since a somewhat more general theorem will be proved later.

5. **Two metrizability conditions.** The requirement in the preceding theorem that there exist a bounded open set in the domain is a considerable restriction as the following theorem shows.

THEOREM 4. *A linear topological space is metrizable as an F-space if there exists a bounded open set in the space.*

Suppose G is a bounded open set. One may assume $\theta \in G$ for otherwise one would simply consider the set $G_x = G - x$, where $x \in G$, also a bounded open set. First note that for any real α the set αG is open. Next, that the set of open sets $\{n^{-1}G\}$ is equivalent to $\mathfrak{U} = \{U\}$. For, since $n^{-1}G$ is open, one can find for any n a neighborhood $U(n) \in \mathfrak{U}$ such that $U(n) \subset n^{-1}G$. On the other hand, let $U \in \mathfrak{U}$ be arbitrary. Then, by Axiom 3 of von Neumann, there exists $V \in \mathfrak{U}$ such that $\beta V \subset U$ for $|\beta| \leqq 1$. Since G is bounded there exists α such that $G \subset \alpha V$. Now select some integer $n \geqq |\alpha|$ so that $|n^{-1}\alpha| \leqq 1$. Then

[4] The definition is used by Mazur and Orlicz for F-spaces. Kolmogoroff [3, p. 30] has used it for topological spaces. It is originally due to Banach.

[5] This lemma has been stated by D. H. Hyers, Bull. Amer. Math. Soc., vol. 43 (1937), p. 203, Abstract 228.

$n^{-1}G \subset n^{-1}\alpha V \subset U$. This proves the equivalence. The space then satisfies the first denumerability axiom of Hausdorff. One may now use the result of Garrett Birkhoff [**8**, p. 428] that a Hausdorff group is metrizable if and only if it satisfies the first denumerability axiom. Since a linear space satisfying the axioms of von Neumann is a Hausdorff group with respect to addition, the result applies here also. To show that the metric is an F-metric one must show (1) $\rho(x, y) = \rho(x - y, \theta)$, (2) $\alpha_n \to 0$, $x \in L$ imply $\alpha_n x \to \theta$, (3) α real, $x_n \to \theta$ imply $\alpha x_n \to \theta$. (1) follows immediately from the fashion in which Birkhoff defines his metric. One may show (2) and (3) independently of metrizability considerations. In fact, (2) and (3) are immediate consequences of the continuity of multiplication. This completes the proof.[6]

The converse of this theorem is not true, i.e., there exist F-spaces in which no sphere is bounded. This is true, for example, of the spaces (S) and (s) [**1**, pp. 9, 10].

In the last part of the proof of the preceding theorem the following theorem was also proved.

THEOREM 5. *A linear topological space is metrizable as an F-space if and only if it satisfies the first denumerability axiom.*[7]

From this theorem and from Kolmogoroff's axioms for a linear topological space the following corollary is easily obtained.

COROLLARY. *Any linear metric space in which the operations of multiplication and addition are continuous may be metrized with an equivalent F-metric.*

In case the space is convex, i.e., satisfies Axiom (6), one may prove the following theorem.

THEOREM 6. *A necessary and sufficient condition that a convex linear topological space have an equivalent norm-metric is that it contain a bounded open set.*

If G is a bounded open set, one may assume as in the preceding theorem that $\theta \in G$. Following the notation of von Neumann, let

$$G_{\text{conv}} \equiv E_x\left[x = \sum_{i=1}^{n} \alpha_i x_i, \qquad \alpha_i \geqq 0, \qquad \sum_{i=1}^{n} \alpha_i = 1, \qquad x_i \in G\right].$$

LEMMA 1. *If L is a convex space, the boundedness of G is both necessary and sufficient for the boundedness of G_{conv}.*

This has been proved by von Neumann [**4**, p. 10, Theorem 14].

LEMMA 2. *If G is an open set, then also G_{conv} is open.*

Let $x = \sum_{k=1}^{n} \alpha_k x_k$ be any element of G_{conv} where $x_k \in G$. Since G is open, there exist neighborhoods U_k such that $x_k + U_k \subset G$. From Axiom (2) one easily

[6] An F-space is generally required to be also complete. This has not been done here. However, sequential completeness, a notion which will be defined later, is sufficient to obtain completeness in the metric sense.

[7] This theorem has been stated by D. H. Hyers, loc. cit.

deduces the existence of $U \in \mathfrak{U}$ such that $U \subset \mathfrak{P}(U_1, \cdots, U_n)$. Then $x_k + U \subset G$ for $k = 1, \cdots, n$. Hence

$$\alpha_1(x_1 + U) + \cdots + \alpha_n(x_n + U) = \alpha_1 x_1 + \cdots + \alpha_n x_n + \alpha_1 U + \cdots + \alpha_n U \subset G_{\text{conv}}.$$

Since $\alpha_1 U + \cdots + \alpha_n U \supset (\alpha_1 + \cdots + \alpha_n)U = U$, one sees that $\alpha_1 x_1 + \cdots + \alpha_n x_n + U \subset G_{\text{conv}}$, i.e., G_{conv} is open.

It now follows from these two lemmas that G_{conv} is a bounded, open, convex set. The following theorem of Kolmogoroff's [**3**, p. 30] may now be applied to obtain the theorem to be proved: *A topological linear space has an equivalent norm topology if and only if there exists a bounded open convex set in the space.*

6. **Extension of Theorem** 3. Theorem 4 shows that the conditions for which Theorem 3 has been stated are rather stringent since it allows for the domain of the transformation $T(x)$ only a proper subclass of the class of F-spaces. This is too restrictive, for wherever the domain is any F-space and the range is any linear topological space, it is easy to obtain the converse of Theorem 2. The method of proof is the same as that used by Mazur and Orlicz [**6**, p. 153].

THEOREM 3′. *If L_1 is an F-space and L_2 a linear topological space, an additive transformation $T(x)$ is continuous if it takes bounded sets into bounded sets.*

Suppose $x_n \to \theta$, i.e., $| x_n | = \rho(x_n, \theta) \to 0$. Then there exists a sequence of integers $k_n \to \infty$ such that $k_n | x_n | \to 0$. Since $k_n | x_n | \geqq | k_n x_n |$, $k_n x_n \to \theta$ and therefore $\{k_n x_n\}$ is bounded. Hence $\{T(k_n x_n)\}$ is bounded, so that $(k_n)^{-1} T(k_n x_n) = T(x_n) \to \theta$. Now suppose there exists $V_0 \in \mathfrak{B}$ such that for each sphere $S_n = E_x[| x | < n^{-1}]$ there exists $x_n \in S_n$ such that $T(x_n)$ non-$\in V_0$. Then $| x_n | \to 0$, i.e., $x_n \to \theta$. Hence $T(x_n) \to \theta$ so that there must exist n_0 such that, for $n \geqq n_0$, $T(x_n) \in V_0$. This is a contradiction.

7. **Strongly bounded transformations**. An additive transformation $T(x)$ will be called strongly bounded if for every $U \in \mathfrak{U}$ and $V \in \mathfrak{B}$ it is possible to find α such that $T(U) \subset \alpha V$, i.e., for each $U \in \mathfrak{U}$, $T(U)$ is a bounded set. The following theorem is then true.

THEOREM 7. *An additive transformation $T(x)$ which is strongly bounded is continuous.*

Let $V \in \mathfrak{B}$ be arbitrary; also $U \in \mathfrak{U}$. By Axiom 3 there exists $V_1 \in \mathfrak{B}$ such that $\alpha V_1 \subset V$ for $| \alpha | \leqq 1$. By hypothesis there exists β such that $T(U) \subset \beta V_1$. Let k be any integer such that $| k^{-1}\beta | \leqq 1$. Then there exists $U_1 \in \mathfrak{U}$ such that $kU_1 \subset U$. Hence $T(kU_1) \subset T(U) \subset \beta V_1$, or $T(U_1) \subset k^{-1}\beta V_1 \subset V$, i.e., $T(x)$ is continuous.

The converse of this theorem is true provided the neighborhoods in the domain are bounded sets. This would imply, however, that the domain is metrizable as an F-space in which spheres are bounded sets.

8. **Linear transformations in convex spaces.** In a convex space pseudo-norms have been defined by von Neumann [**4**, pp. 18, 19] as follows. Let $\| x \|_U^+$ equal the greatest lower bound of $\alpha > 0$ such that $x \in \alpha U$ and let $\| x \|_U = \max (\| x \|_U^+, \| -x \|_U^+)$. He has shown that $\| x \|_U$ has the following properties: $\| x + y \|_U \leqq \| x \|_U + \| y \|_U$, $\| \alpha x \|_U = |\alpha| \cdot \| x \|_U$, $\| \theta \|_U = 0$, $\| x \|_U$ is a continuous function of x. The pseudo-norm $\| x \|_U$ lacks the norm property $\| x \| = 0$ if and only if $x = \theta$. However, by use of pseudo-norms, one may obtain generalizations of several well known theorems for normed spaces. An existence theorem will be proved first.

THEOREM 8. *If L is a convex space, then for any $x_0 \in L$ and any $U \in \mathfrak{U}$ there exists a linear continuous functional $F(x)$ defined on L with the property that $F(x_0) = \| x_0 \|_U$.*

Denote by K the linear set $E_x[\alpha x_0, \alpha \text{ real}]$ and define on K the functional $f(\alpha x_0) = \alpha \| x_0 \|_U$. Then $f(x)$ is obviously linear on K. Then, if we denote $\| x \|_U$ by $p(x)$, it follows that $p(x + y) \leqq p(x) + p(y)$, $p(tx) = tp(x)$ for $t \geqq 0$ and $p(x) \geqq f(x)$ for $x \in K$. One may now use a theorem of Banach's on the extension of additive functionals [**1**, p. 27, Theorem 1] and assert that there exists a linear functional $F(x)$ defined on all L such that $F(x) = f(x)$ for $x \in K$ and $-\| x \|_U \leqq F(x) \leqq \| x \|_U$ for all x. Since $\| x \|_U$ is continuous at θ, it follows that $F(x)$ is also continuous at θ and hence, since it is additive, that it is continuous.

By the use of pseudo-norms a bounded, additive transformation may be defined in a manner analogous to the definition in a normed space. An additive transformation $T(x)$ on L_1 to L_2, L_1 and L_2 convex spaces, will be said to be *bounded* if to each $V \in \mathfrak{V}$ correspond $U \in \mathfrak{U}$ and a real number $M(U, V) \geqq 0$ such that $\| T(x) \|_V \leqq M(U, V) \| x \|_U$. Such a transformation obviously takes bounded sets into bounded sets. The following theorem connecting bounded and continuous additive transformations is proved.

THEOREM 9. *An additive transformation $T(x)$ is continuous if and only if it is bounded.*

If $T(x)$ is continuous, for each $V \in \mathfrak{V}$ one may find $U \in \mathfrak{U}$ such that $T(U) \subset V$, or, using pseudo-norms, such that $\| x \|_U < 1$ implies $\| T(x) \| < 1$. Note that if $\| x \|_U = 0$, then $\| T(x) \|_V = 0$ for otherwise one could find a real α such that $\| T(\alpha x) \|_V = |\alpha| \cdot \| T(x) \|_V > 1$ even though $\| \alpha x \|_U = 0$. Suppose now that the required $M(U, V)$ does not exist. Select a sequence of real positive integers $M_n \to \infty$ and assume that to each n corresponds an x_n such that $\| T(x_n) \|_V > M_n \| x_n \|_U \neq 0$. Let

$$y_n = \frac{x_n}{M_n \| x_n \|_U}.$$

Then

$$\| T(y_n) \|_V = \frac{1}{M_n \| x_n \|_U} \| T(x_n) \|_V > 1.$$

But $\| y_n \|_U = M_n^{-1} \to 0$, so that one can find n_0 such that, for $n \geqq n_0$, $\| y_n \|_U < 1$ so that $\| T(y_n) \|_V < 1$, a contradiction for $n \geqq n_0$.

On the other hand, suppose $T(x)$ is bounded. Then for arbitrary $V \epsilon \mathfrak{V}$ one may find $U \epsilon \mathfrak{U}$ and $M(U, V) \geqq 0$ such that $\| T(x) \|_V \leqq M(U, V) \cdot \| x \|_U$. Then $M(U, V) \| x \|_U < 1$ or $x \epsilon [M(U, V)]^{-1}U$ implies that $T(x) \epsilon V$. Since there exists $U_1 \subset [M(U, V)]^{-1}U$, it follows that $T(U_1) \subset V$, i.e., $T(x)$ is continuous.

The following corollary is an immediate consequence.

Corollary. *If L is convex and if $F(x)$ is an additive functional defined on L, $F(x)$ is continuous if and only if there exist $U \epsilon \mathfrak{U}$ and $M(U) \geqq 0$ such that $| F(x) | \leqq M(U) \| x \|_U$.*

In connection with this corollary the following theorem may be proved.

Theorem 10. *Suppose a linear functional $F(x)$ on a convex space L is continuous if and only if for every $U \epsilon \mathfrak{U}$ there exists $M(U) \geqq 0$ such that $| F(x) | \leqq M(U) \| x \|_U$. Then any of the pseudo-norms $\| x \|_U$ for $U \epsilon \mathfrak{U}$ is also a norm and the corresponding norm topology is not stronger[8] than the given topology but is equivalent to it with respect to the determination of the class of linear and continuous functionals.*

Since pseudo-norms have all the norm properties except $\| x \|_U = 0$ implies $x = \theta$, only this need be shown. Suppose $\| x_0 \|_U = 0$ for some $U \epsilon \mathfrak{U}$. First note that there exists $U_0 \epsilon \mathfrak{U}$ such that $\| x_0 \|_{U_0} \neq 0$ unless $x_0 = \theta$. Then, by Theorem 8, there exists a linear continuous functional $F_0(x)$ such that $F_0(x_0) = \| x_0 \|_{U_0} \neq 0$. On the other hand, since $| F(x) | \leqq M(U, F) \| x \|_U$, $F(x_0) = 0$ for all linear continuous F. Hence $F_0(x_0) = 0$ which is a contradiction unless $x_0 = \theta$. Therefore $\| x \|_U = 0$ if and only if $x = \theta$, i.e., $\| x \|_U$ is a true norm. If one now fixes $U_0 \epsilon \mathfrak{U}$ one may define the "spherical" neighborhoods in the corresponding norm topology, viz., $S_\alpha = E_x[\| x \|_{U_0} < \alpha]$. From the fact that for any α there exists U_1 such that $U_1 \subset \alpha U_0$ it follows that any sphere S_α contains a neighborhood of $\mathfrak{U}$, i.e., the norm topology is not stronger than the given topology. From the hypothesis of the theorem any linear functional which is continuous in the neighborhood topology is continuous in the norm topology. On the other hand, since any linear functional which is continuous in the norm topology is bounded, i.e., there exists $M \geqq 0$ such that $| F(x) | \leqq M \| x \|_U$, it follows from the corollary to Theorem 9 that $F(x)$ is continuous in the neighborhood topology.

The following theorems are generalizations of well known theorems for normed spaces [**1**, p. 55, Theorem 4; p. 57, Lemma; p. 58, Theorem 6]. The proofs will not be given since they are quite similar to those for the corresponding theorems in normed spaces.

Theorem 11. *If $f(x)$ is any functional defined on G, a subset of a convex*

[8] One topology will be said to be not stronger than another if every limit point in the sense of the first is also a limit point in the sense of the second.

space L, there exists a linear continuous functional $F(x)$ defined on all L such that

(1) $F(x) = f(x)$ *for* $x \in G$,

(2) $|F(x)| \leqq M \| x \|_U$ *for given* $M \geqq 0$ *and* $U \in \mathfrak{U}$ *if and only if for all finite sets* $\{x_i\} \subset G$ *and real numbers* $\{\alpha_i\}$

$$\left| \sum_{i=1}^{n} \alpha_i f(x_i) \right| \leqq M \cdot \left\| \sum_{i=1}^{n} \alpha_i x_i \right\|_U .$$

THEOREM 12. *If $G \subset L$ is a linear set, $U \in \mathfrak{U}$ any neighborhood, and $y_0 \in L$ an element such that, for $x \in G$, $\| y_0 - x \|_U \geqq d > 0$, there exists a linear continuous functional $F(x)$ defined on L such that*

(1) $F(y_0) = 1$,

(2) $F(G) = 0$,

(3) $|F(x)| \leqq d^{-1} \| x \|_U$. *If d is the greatest lower bound of the values* $\| y_0 - x \|_U$ *for* $x \in G$, *then in (3)* d^{-1} *is the smallest value M such that* $|F(x)| \leqq M \| x \|_U$.

THEOREM 13. *If $G \subset L$ is an arbitrary set and if $y_0 \in L$, a necessary and sufficient condition that y_0 belong to the closed linear extension of G is that $F(x) = 0$ for $x \in G$ imply $F(y_0) = 0$ for all linear continuous $F(x)$.*

9. **On the domain of a linear continuous functional.** In order to provide a motivation for and examples of the succeeding theorem the general form of the linear continuous functional on three F-spaces is given.

Space (s). The general form of the linear continuous functional on (s) is $F(x) = \sum_{i=1}^{n} \varphi_i \xi_i$, where n and $\varphi_1, \cdots, \varphi_n$ depend on F and where $x = (\xi_1, \xi_2, \cdots)$ [**1**, p. 50, Theorem 11].

Space (S). The most general linear functional on (S) is the zero-functional [**1**, p. 234].

Denote by (CF) the set of all continuous functions defined on $[-\infty, \infty]$ with the following metric: $\rho(x, y) = \sum_{n=1}^{\infty} \frac{(x, y)_n}{1 + (x, y)_n}$, where $(x, y)_n = \max_{|t| \leqq n} |x(t) - y(t)|$ [cf. **9**, pp. 30, 31].

Space (CF). Every linear continuous functional on (CF) is of the form $F(x) = \int_{\alpha}^{\beta} x(t)\, d\varphi(t)$, where α, β and $\varphi(t)$ depend on F and $\varphi(t)$ is of bounded variation in the interval $[\alpha, \beta]$.

Since the proof of this involves no unusual procedure, it is not given here.

In each of these examples the functional seems to operate essentially only on a "part" of each element of the space in such a manner that when a set of elements converges according to the F-metric, their corresponding "parts" converge according to a norm-metric. The precise meaning of this is given by the next theorem. Before stating this theorem a definition will be given.

DEFINITION. A set $X \subset L$ will be called a *flat subspace* if for arbitrary $x \in X$, $X - x$ forms a linear set.

THEOREM 14. *If $F(x)$ is a linear continuous functional on a linear topological space L, there corresponds to each $x \epsilon L$ a flat subspace X such that*

(1) *if $x_1, x_2 \epsilon X$, then $F(x_1) = F(x_2) \equiv F(X)$,*

(2) *there exists a norm topology, $\| X \|_F$, for the set, L_F, of such subspaces such that if $x \epsilon \bar{R}$, $X \epsilon \bar{R}_F$,*

(3) *if $\| X_n - X \|_F \to 0$, then $F(X_n) \to F(X)$,*

(4) *the linear continuous functionals on L_F determined by $\| X \|_F$ are linear continuous functionals on L with $F(x) \equiv F(X)$ for $x \epsilon X$.*

Define $X \equiv E_x[x = \bar{x} + x'$, where $F(x) = F(\bar{x})]$. Each $x \epsilon L$ is obviously in some X. Letting $\Theta = E_x[F(x) = 0]$, we easily see that Θ is a linear set. Since $X = \Theta + \bar{x}$, it follows that X is a flat subspace. It also easily follows from the linearity of L and of F that L_F is a linear set.

(1) is an obvious consequence of the definition of X.

Consider (2). A topology will now be defined for L_F. If R is a set in L, the set of elements in L_F corresponding to the elements of R will be denoted by R_F. The closure of a set R_F will be defined as follows: $X \epsilon \bar{R}_F$ if $F(X) \epsilon \overline{F(R_F)}$. Although $x \epsilon \bar{R}$ does not imply that x is a sequential limit point of R, the topology in L_F will obviously be of a sequential type. This topology also satisfies Kolmogoroff's axioms for a linear topological space, the continuity of addition and multiplication following from the linearity and continuity of $F(x)$. Now let $S_\alpha = E_X[| F(X) | < \alpha]$. It then follows from the linearity of $F(x)$ that S_α is a bounded, open, and convex subset of L_F. The normability condition of Kolmogoroff used in the proof of Theorem 6 may now be applied to give an equivalent norm topology for L which will be denoted by $\| X \|_F$. Then, if $x_0 \epsilon \bar{R}$, it follows from continuity of $F(x)$ that $F(x_0) \epsilon \overline{F(R)}$ and therefore $F(X_0) \epsilon \overline{F(R_F)}$, i.e., $X_0 \epsilon \bar{R}_F$.

(3) follows immediately from the definition of $\| X \|_F$.

In order to prove (4) let $\mathfrak{F}_0$ be the set of all linear continuous functionals on L_{F_0} determined by $\| X \|_{F_0}$. If $F \epsilon \mathfrak{F}_0$, define $F(x) = F(X)$ for $x \epsilon X$. Then, if $x_0 \epsilon \bar{R}$, $X_0 \epsilon \bar{R}_{F_0}$ by (2). By continuity of $F(X)$, $F(X_0) \epsilon \overline{F(R_{F_0})}$ and hence $F(x_0) \epsilon \overline{F(R)}$, i.e., $F(x)$ is continuous.

10. **Category of linear topological spaces**. In the study of complete metric spaces one of the fundamental theorems as far as the theory of functions for such spaces is concerned is the one which states that every such space is of the second category. Since this property seems to be important in the proofs of many theorems, one should like a definition of completeness for topological linear spaces to give the same or a corresponding theorem. J. von Neumann has shown how the ordinary notion of "sequential" completeness can be carried over into such spaces and, then, in order to avoid a certain difficulty in connection with it, proposes to call a topological linear space "topologically" complete if every closed and totally bounded set is compact [**4**, pp. 1–3]. The following theorem shows that for neither of these definitions the category theorem holds.

In order to state the theorem the weak neighborhood topology for a normed space L must be defined. For each finite set of linear continuous functionals, $F_1, \cdots, F_n$, defined on L and for each $\delta > 0$, let the neighborhood

$$U(F_1, \cdots, F_n; \delta) \equiv E_x[|F_i(x)| < \delta, i = 1, \cdots, n].$$

The set of all such neighborhoods will be denoted by $\mathfrak{U}$ and defines the weak topology[9] for L. These neighborhoods satisfy Axioms (1)–(6), i.e., the space is a convex linear topological space. The theorem is then the following.

THEOREM 15. *A normed space with its weak neighborhood topology for which there exists no finite total set of functionals is of the first category.*

The proof will be broken up into several lemmas.

LEMMA 1. *If no neighborhood of a topological linear space L is bounded, every bounded set is non-dense.*

If a set $M \subset L$ is bounded, then also its closure $\bar{M}$ is bounded [**4**, p. 9, Theorem 11]. Since no neighborhood is bounded, it would be impossible to find for any $x \in \bar{M}$ a $U \in \mathfrak{U}$ such that $x + U \subset \bar{M}$, i.e., $\bar{M}$ has a vacuous interior. But this is just the condition that M be non-dense [**10**, p. 31, II].

LEMMA 2. *The set $S_n = E_x[\| x \| \leqq n]$ is bounded.*

Consider arbitrary $U(F_1, \cdots, F_k; \delta) \in \mathfrak{U}$. Let $m = \max(\| F_1 \|, \cdots, \| F_k \|)$ and let $\alpha = 2mn\delta^{-1}$. Then $\alpha U(F_1, \cdots, F_k; \delta) = E_x[|F_i(x)| < \alpha\delta = 2mn]$. Now suppose $x \in S_n$. Then $|F_i(x)| \leqq \| F_i \| \cdot \| x \| \leqq mn < 2mn$, i.e., $x \in \alpha U$. Therefore $S_n \subset \alpha U(F_1, \cdots, F_k; \delta)$ and S_n is bounded since U was arbitrary.

LEMMA 3. *If L does not have a finite total set of functionals, then none of its weak neighborhoods are bounded.*

Consider arbitrary $U(F_1, \cdots, F_n; \delta) \in \mathfrak{U}$. Let $G = E_x[F_i(x) = 0, i = 1, \cdots, n]$. Since the set, $F_1, \cdots, F_n$, cannot be total, G must contain other elements than θ. Then, since G is a linear set, one can find $x_0 \in G$ with $\| x_0 \| = 1$. Also, by a theorem of Banach's [**1**, p. 55, Theorem 3], there exists a linear continuous functional $F_0(x)$ such that $F_0(x_0) = \| x_0 \| = 1$. Now consider the neighborhood $U(F_0; 1) = E_x[|F_0(x)| < 1]$ and the sequence $x_0, 2x_0, \cdots, kx_0, \cdots$. Since $kx_0 \in G$, every element of the sequence is contained in $U(F_1, \cdots, F_n; \delta)$. However, since $mU(F_0; 1) = E_x[|F_0(x)| < m]$, one easily sees that, for $k \geqq m$, kx_0 non-$\in$ $mU(F_0; 1)$. It is then obvious that there exists no number α such that $U(F_1, \cdots, F_n; \delta) \subset \alpha U(F_0; 1)$, i.e., $U(F_1, \cdots, F_n; \delta)$ which was arbitrary is not bounded.

From the three lemmas it follows that the set S_r is non-dense. Then, since $L = \sum_{n=1}^{\infty} S_n$, it is of the first category in itself.

Not all "weak" normed spaces are sequentially complete, an important

[9] Note that the weak neighborhood topology is not necessarily equivalent to the weak sequential topology for L (i.e., where $x_k \to x$ is defined as meaning $F(x_k) \to F(x)$ for all linear continuous F). In fact, in Hilbert space there exist points which are limit points of sets without being sequential limit points [**5**, p. 380]. The two topologies are equivalent, however, as far as convergence of sequences is concerned.

example being the space (C) of all continuous functions defined on the interval $[0, 1]$. However, the spaces L_p and l_p for $p \geqq 1$ are sequentially complete in their weak topologies [**1**, pp. 140–143, §4].[10] J. von Neumann [**4**, p. 17, Theorem 23] has shown that abstract Hilbert space with its weak neighborhood topology is also topologically complete. Therefore neither definition is sufficient to obtain the property that the space be of the second category. The additional remark should be made that none of the spaces L_p, l_p or abstract Hilbert space have finite total sets of functionals defined on them.

The extent of the restriction in the preceding theorem that the normed space have no finite total set of functionals is given by the following theorem.

THEOREM 16. *A normed space L has a finite total set of functionals defined on it if and only if it is finite dimensional.*

Suppose $F_1, \cdots, F_n$ is a total set of functionals for L and assume that for every k there exist $z_1, \cdots, z_k \subset L$ such that $\sum_{i=1}^{k} \alpha_i z_i = \theta$ implies $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Now consider the set of equations

$$\sum_{j=1}^{m} \beta_j F_i(z_j) = 0 \qquad (i = 1, \cdots, n),$$

where the β_j are to be the unknowns. Then, if $m > n$, it is obvious that the rank of the matrix of the coefficients, $(F_i(z_j))$, cannot exceed n. By a well known theorem on solutions of linear homogeneous equations, there exist at least $m - n > 0$ linearly independent solutions of the set of equations, i.e., there exists $z = \sum_{j=1}^{m} \beta_j z_j \neq \theta$ such that $F_i\left(\sum_{j=1}^{m} \beta_j z_j\right) = 0$ for $i = 1, \cdots, n$. But, by the totality of the F_i, one must have $\sum_{j=1}^{m} \beta_j z_j = \theta$ which is a contradiction. Therefore the space is at most n-dimensional.

Suppose L is n-dimensional, i.e., there exist $z_1, \cdots, z_n$, linearly independent, such that to each $x \in L$ there corresponds a set of real numbers $\alpha_1, \cdots, \alpha_n$ with the property that $x = \sum_{i=1}^{n} \alpha_i z_i$. Let G_i be the linear manifold determined by $z_1, \cdots, z_{i-1}, z_{i+1}, \cdots, z_n$. Then there exists a linear continuous functional $F_i(x)$ such that $F_i(z_i) = 1$ and $F_i(G_i) = 0$ [**1**, p. 57, Lemma]. Then $F_1, \cdots, F_n$ is a total set. For suppose that, for $x = \sum_{j=1}^{n} \alpha_j z_j$, $F_i(x) = 0$ $(i = 1, \cdots, n)$. Then $F_i\left(\sum_{j=1}^{n} \alpha_j z_j\right) = \sum_{j=1}^{n} \alpha_j F_i(z_j) = \alpha_i = 0$, i.e., $x = \theta$.

One may now obtain some additional information concerning the metrizability of the weak neighborhood topologies, viz., the following theorem.

[10] Since this is a sequential notion it makes no difference that the "sequential" and "neighborhood" weak topologies are not equivalent as long as the "sequential part" of the latter is equivalent to the former.

THEOREM 17. *The weak neighborhood topology is metrizable as a norm-metric if and only if the space is finite dimensional.*

The non-metrizability of the infinite dimensional spaces follows from Lemma 3 and Theorem 6. That the finite dimensional spaces may be normed follows from Tychonoff's result [**6**, p. 769] that every finite dimensional linear topological space is homeomorphic to Euclidean space of the same dimension.

The situation with regard to metrizability of weak neighborhood topologies is different from that in the weak sequential topologies, for it is well known that in the space l_1 the weak sequential topology is equivalent to the norm topology [cf. **1**, p. 139]. One easily concludes then that in the weak neighborhood topology for l_1 there must exist points which are limit points without being sequential limit points.[11]

It is interesting to note that Hilbert space is also of the first category with its weak sequential topology. Let $\{\varphi_n\}$ be an orthonormal set. Then, since the Fourier coefficients $(f, \varphi_n) \to 0$ for all $f \in \mathfrak{H}$, φ_n converges weakly to θ and $3r\varphi_n + f$ converges weakly to f. But, if $S_r = E_f[\| f \| \leqq r]$, $3r\varphi_n + f$ non-$\in S_r$ if $f \in S_r$ so that each element of S_r is a limit point of a sequence not in S_r, i.e., S_r is non-dense. The rest of the proof follows as in Theorem 15.

11. **Linear transformations in weak normed spaces.** In this section we return to the discussion of the relation between bounded sets and continuity of additive transformations. The theorem which will be proved has some interest in itself and also serves as an example to show that Theorem 3′ may hold even when the domain is not metrizable.

THEOREM 18. *If L_1 and L_2 are two normed spaces with their weak neighborhood topologies, any additive transformation $T(x)$ which takes bounded sets into bounded sets is continuous.*

If $R \subset L_1$ is bounded in the weak topology, there must exist for each linear continuous functional, $F(x)$, on L_1 a number $M(F, R) \geqq 0$ such that $| F(R) | \leqq M(F, R)$, for otherwise for some $F(x)$ on L_1 there would exist $x_n \in R$ and $\alpha_n \to 0$ such that $F(\alpha_n x_n)$ does not converge to zero. First it will be shown that if $\| x_n \| \to 0$, then $T(x_n) \to \theta$ weakly. For if $\| x_n \| \to 0$ there exist positive integers $k_n \to \infty$ such that $\| k_n x_n \| = k_n \| x_n \| \to 0$. Since $\| k_n x_n \| \to 0$, $k_n x_n \to \theta$ weakly and hence $\{k_n x_n\}$ is weakly bounded. From the hypothesis it follows that $\{T(k_n x_n)\}$ is also weakly bounded. By the remark above there will exist for each linear continuous F an $M(F) \geqq 0$ such that $| F(T(k_n x_n)) | \leqq M(F)$ or, using additivity of F and T, such that $| F(T(x_n)) | \leqq k_n^{-1} M(F) \to 0$, i.e., $T(x_n) \to \theta$ weakly. Now, from the fact that $\| x_n \| \to 0$ implies $T(x_n) \to \theta$ weakly, one may show that $\| x_n \| \to 0$ implies $\| T(x_n) \| \to 0$, i.e., $T(x)$ is continuous in the norm topologies. This will be proved by showing that there

[11] A direct construction of such a point has been given by Köthe [**11**, p. 118]. Let $R = E_x[\| x \| \geqq 1]$. Then θ can obviously not be a sequential limit point of R, but, since $L - R$ is bounded, $\mathfrak{P}(R, U)$ is not vacuous for any $U \in \mathfrak{U}$. Hence θ is a limit point of R

exists $M \geqq 0$ such that $\| T(x) \| \leqq M \| x \|$. Suppose there exists no such M. Then for each integer m there exists $y_m \in L_1$ such that $\| T(y_m) \| \geqq m^2 \| y_m \|$. Let $z_m = \dfrac{y_m}{m \| y_m \|}$. Then $\| z_m \| \to 0$ and hence $T(z_m) \to \theta$ weakly. However,

$$\| T(z_m) \| = \frac{1}{m \| y_m \|} \| T(y_m) \| > m,$$

i.e., the set $\{T(z_m)\}$ is not of bounded norm. This is contradictory, for it is well known that a sequence can converge weakly to θ only if it is of bounded norm [**1**, p. 133, Theorem 1]. Hence the transformation $T(x)$ must be a bounded transformation and therefore continuous in the norm topologies. But then $T(x)$ must also be continuous in the weak neighborhood topologies. For if F is a linear continuous functional on L_2, then $F(T)$ is a linear continuous functional on L_1. Hence, if $V(F_1, \cdots, F_n; \delta)$ is an arbitrary neighborhood in L_2, $U(F_1(T), \cdots, F_n(T); \delta)$ is a neighborhood in L_1 and has the property that $T(U) \subset V$, i.e., $T(x)$ is continuous in the weak neighborhood topologies.

Bibliography

1. S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne, Tom I, Warszawa, 1932.
2. M. Fréchet, *Les Espaces Abstraits*, Paris, 1928.
3. A. Kolmogoroff, *Zur Normierbarkeit eines allgemeinen topologischen linearen Raumes*, Studia Math., vol. 5 (1934), pp. 29–33.
4. John von Neumann, *On complete topological spaces*, Trans. Amer. Math. Soc., vol. 37 (1935), pp. 1–20.
5. John von Neumann, *Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren*, Math. Ann., vol. 102 (1929), pp. 370–427.
6. A. Tychonoff, *Ein Fixpunktsatz*, Math. Ann., vol. 111 (1935), pp. 767–776.
7. S. Mazur and W. Orlicz, *Über Folgen linearen Operationen*, Studia Math., vol. 4 (1933), pp. 152–157.
8. Garrett Birkhoff, *A note on topological groups*, Compositio Math., vol. 3 (1936), pp. 427–430.
9. R. Gateaux, *Sur diverses questions du calcul fonctionnel*, Bull. Soc. Math. France, vol. 40 (1922), pp. 1–37.
10. C. Kuratowski, *Topologie* I, Monografje Matematyczne, Tom III, Warszawa, 1933.
11. G. Köthe, *Die Teilräume eines linearen Koordinatenraumes*, Math. Ann., vol. 114 (1937), pp. 99–125.
12. C. R. Adams, *The space of functions of bounded variation and certain general spaces*, Trans. Amer. Math. Soc., vol. 40 (1936), pp. 421–438.

Brown University.

INTEGRAL FUNCTIONS BOUNDED ON SEQUENCES OF POINTS

By Norman Levinson

1. As an extension of results of V. Ganapathy Iyer we shall prove the following

Theorem. *Let* $\{\lambda_n\}$ *and* $\{\mu_n\}$ *be two increasing sequences such that*

$$\lim_{n\to\infty} \frac{n}{\lambda_n} = D_\lambda, \qquad \lim_{n\to\infty} \frac{n}{\mu_n} = D_\mu.$$

Let the indices of condensation[1] *of these sequences be zero. Let* $f(z)$ *be an integral function such that*

$$f(\pm\lambda_n) = O(1), \qquad f(\pm i\mu_n) = O(1), \tag{1.0}$$

and

$$\overline{\lim_{|z|\to\infty}} \frac{\log|f(z)|}{|z|} = k < \pi(D_\lambda^2 + D_\mu^2)^{\frac{1}{2}}. \tag{1.1}$$

Then $f(z)$ *is a constant.*

That the above theorem is a best possible result follows from trivial considerations (for example, $f(z) = \sin \pi D_\lambda z \sinh \pi D_\mu z$). Iyer[2] has proved that the above result holds under more stringent conditions; namely, with (1.1) replaced by

$$k < \pi \min(D_\lambda, D_\mu), \tag{1.2}$$

or else with (1.0) replaced by

$$\lim_{n\to\infty} \frac{\log|f(\pm\lambda_n)|}{\lambda_n} = \lim_{n\to\infty} \frac{\log|f(\pm i\mu_n)|}{\mu_n} = -\infty.$$

The method of proof used here is an extension of a method used in proving certain simpler theorems.[3]

2. The proof of our theorem will be given in two parts. Part 1 is the essential part.

Received November 12, 1937.

[1] The term index of condensation is defined in *Séries de Dirichlet*, Vladimir Bernstein, Paris, 1933, p. 25. $\underline{\lim}_{n\to\infty} (\lambda_{n+1} - \lambda_n) > 0$ is sufficient for zero index of condensation.

[2] V. Ganapathy Iyer, *On the order and type of integral functions bounded at a sequence of points*, Annals of Mathematics, vol. 38 (1937), p. 311.

[3] N. Levinson, *On certain theorems of Pólya and Bernstein*, Bulletin of the American Mathematical Society, vol. 42 (1936), p. 702. The method is an extension of that used in proving Theorem 2, p. 703.

Let

$$A(z) = \prod_1^\infty \left(1 - \frac{z^2}{\lambda_n^2}\right), \qquad B(z) = \prod_1^\infty \left(1 + \frac{z^2}{\mu_n^2}\right).$$

Let

$$a = \pi D_\lambda, \qquad b = \pi D_\mu, \qquad \tan \alpha = \frac{a}{b}. \tag{2.0}$$

$A(z)$ possesses the following well-known properties.[4] For any $\epsilon > 0$

$$\frac{1}{A'(\pm\lambda_n)} = O(e^{\epsilon\lambda_n}), \tag{2.1}$$

$$A(re^{i\theta}) = O(e^{ar|\sin\theta| + \epsilon r}), \tag{2.2}$$

$$\frac{1}{A(re^{i\theta})} = O(e^{-ar|\sin\theta| + \epsilon r}), \tag{2.3}$$

where in (2.3) θ is not an integral multiple of π. Similar results hold for $B(z)$. Also there exists[5] an increasing sequence of positive numbers $\{R_n\}$ such that $R_n \to \infty$ as $n \to \infty$, and for any $\epsilon > 0$

$$\frac{1}{A(R_n e^{i\theta})} = O(e^{\epsilon R_n}), \qquad \frac{1}{B(R_n e^{i\theta})} = O(e^{\epsilon R_n}). \tag{2.4}$$

Also no R_n is nearer than a fixed amount to any λ_m or μ_m.

Proof of Theorem. Part 1. *In this part it is shown that for any* $\epsilon > 0$

$$f(z) = O(e^{\epsilon|z|}). \tag{2.5}$$

For any $\gamma > 0$ let

$$\begin{aligned}\psi(z) = \frac{f(z)}{A(z)\,B(z)} &- \sum_1^\infty \frac{f(\lambda_n)\exp[(z - \lambda_n)(-b - ib^2/a + \gamma)]}{A'(\lambda_n)\,B(\lambda_n)\,(z - \lambda_n)} \\ &- \sum_1^\infty \frac{f(-i\mu_n)\exp[(z + i\mu_n)(-ia - a^2/b + i\gamma)]}{B'(-i\mu_n)\,A(-i\mu_n)\,(z + i\mu_n)}.\end{aligned} \tag{2.6}$$

The series on the right converge by (2.1) and (2.3). For any $\epsilon > 0$ and small $\delta > 0$,

$$\begin{aligned}\psi(re^{i(\alpha-\delta)}) = O\Bigl(&\exp\{-[(a^2 + b^2)^{\frac{1}{2}}\cos\delta - k - \epsilon]r\} \\ &+ \exp\left\{-(b - \gamma)r\cos(\alpha - \delta) + \frac{b^2}{a} r \sin(\alpha - \delta)\right\} \\ &+ \exp\left\{(a - \gamma)r\sin(\alpha - \delta) - \frac{a^2}{b} r\cos(\alpha - \delta)\right\}\Bigr).\end{aligned} \tag{2.7}$$

[4] V. Bernstein, loc. cit., Note II, p. 267.

[5] Consider the single set $\{\nu_n\}$ made up of $\{\lambda_n\}$ and $\{\mu_n\}$. Let $Q(q, \{\nu_n\})$ be defined as in Bernstein, loc. cit., p. 271. Then the existence of $\{R_n\}$ follows from Note II, no. 4, p. 278.

If we use $\tan \alpha = a/b$, a simple calculation shows

$$-b \cos (\alpha - \delta) + \frac{b^2}{a} \sin (\alpha - \delta) = -(a^2 + b^2)^{\frac{1}{2}} \frac{b}{a} \sin \delta,$$

$$a \sin (\alpha - \delta) - \frac{a^2}{b} \cos (\alpha - \delta) = -(a^2 + b^2)^{\frac{1}{2}} \frac{a}{b} \sin \delta.$$

Thus (2.7) becomes

$$\begin{aligned} \psi(re^{i(\alpha-\delta)}) = O\Big(&\exp\{-[(a^2+b^2)^{\frac{1}{2}} \cos \delta - k - \epsilon]r\} \\ &+ \exp\left\{\gamma r - (a^2+b^2)^{\frac{1}{2}} \frac{b}{a} r \sin \delta\right\} \\ &+ \exp\left\{\gamma r - (a^2+b^2)^{\frac{1}{2}} \frac{a}{b} r \sin \delta\right\}\Big). \end{aligned} \tag{2.8}$$

Let $\epsilon = \frac{1}{2}[(a^2 + b^2)^{\frac{1}{2}} - k]$. Also let

$$p = (a^2 + b^2)^{\frac{1}{2}} \min\left(\frac{a}{b}, \frac{b}{a}\right).$$

Then we can choose $\delta > 0$ so small that $[(a^2 + b^2)^{\frac{1}{2}}\cos \delta - k - \epsilon] > p \sin \delta$. Thus (2.8) becomes

$$\psi(re^{i(\alpha-\delta)}) = O(e^{\gamma r - pr \sin \delta}).$$

We now introduce a small $\epsilon > 0$ and take $\gamma = \epsilon \sin \delta$. Then

$$\psi(re^{i(\alpha-\delta)}) = O(e^{-(p-\epsilon)r \sin \delta}). \tag{2.9}$$

Similarly

$$\psi(re^{i(-\pi+\alpha+\delta)}) = O(e^{-(p-\epsilon)r \sin \delta}). \tag{2.10}$$

By (2.4) and (2.6) for any $\epsilon > 0$,

$$\psi(R_n e^{i\theta}) = O(e^{(k+\epsilon)R_n}) \qquad (-\pi + \alpha < \theta < \alpha).$$

Thus $\psi(z) \exp\{-i(k + \epsilon)ze^{-i\alpha}/\sin \delta\}$ is bounded on the radii am $z = \alpha - \delta$ and am $z = -\pi + \alpha + \delta$ by (2.9) and (2.10), and it is also uniformly bounded on the arcs $|z| = R_n$, $-\pi + \alpha + \delta \leqq \text{am}\, z \leqq \alpha - \delta$. Thus by the maximum modulus theorem it is bounded in the sector $-\pi + \alpha + \delta \leqq \text{am}\, z \leqq \alpha - \delta$. It follows at once that $\psi(z)$ is of exponential type in this sector.

But by well-known results of Pragmén-Lindelöf,[6] $\psi(z)$ of exponential type in the sector, (2.9) and (2.10) give

$$\psi(re^{i\theta}) = O(e^{-(p-\epsilon)r \sin(\alpha-\theta)}) \qquad (-\pi + \alpha + \delta \leqq \theta \leqq \alpha - \delta).$$

[6] For example, see Titchmarsh, *Theory of Functions*, Oxford, 1932, pp. 176-186.

There is no essential difference between the cases $a > b$, $b > a$. Let us assume $a \geqq b$. Then $p = (a^2 + b^2)^{\frac{1}{2}} b/a$ and for $x > 0$

$$\psi(x) = O\left(\exp\left\{-\left[(a^2 + b^2)^{\frac{1}{2}} \frac{b}{a} - \epsilon\right] x \sin \alpha\right\}\right).$$

Thus using (2.2), we see that

$$A(x)\, B(x)\, \psi(x) = O\left(\exp\left\{\left[b + \epsilon - (a^2 + b^2)^{\frac{1}{2}} \frac{b}{a} \sin \alpha\right] x\right\}\right).$$

Since $[(a^2 + b^2)^{\frac{1}{2}} \sin \alpha]/a = 1$,

$$A(x)\, B(x)\, \psi(x) = O(e^{\epsilon x}), \qquad x > 0. \tag{2.11}$$

But

$$f(x) = A(x)\, B(x)\, \psi(x)$$
$$+ B(x) \exp\left[-(b - \gamma)x + ib^2 x/a\right] \sum_1^\infty \frac{f(\lambda_n)\, A(x) \exp\left[\lambda_n(b - \gamma + ib^2/a)\right]}{A'(\lambda_n)\, B(\lambda_n)\, (x - \lambda_n)}$$
$$+ B(x)\, A(x) \exp\left[-(a^2/b - \gamma)x - iax\right] \sum_1^\infty \frac{f(-i\mu_n) \exp\left[\mu_n(a - \gamma - ia^2/b)\right]}{B'(-i\mu_n)\, A(-i\mu_n)\, (x + i\mu_n)}.$$

Thus by (2.2) and (2.11), $f(x) = O(e^{\epsilon x})$, $x > 0$. Similarly this holds for negative x. Thus

$$f(x) = O(e^{\epsilon |x|}).$$

But $f(z)$ is of exponential type. If

$$c = \overline{\lim_{|y| \to \infty}} \frac{\log |f(iy)|}{|y|},$$

then again by Pragmén-Lindelöf

$$f(z) = O(\exp\left[cr \mid \sin \theta \mid + \epsilon r\right]). \tag{2.12}$$

But by a theorem of Bernstein,[7] (2.12) and $\lim_{n \to \infty} n/\mu_n > 0$ imply that

$$c = \overline{\lim_{n \to \infty}} \frac{\log |f(\pm i\mu_n)|}{\mu_n}.$$

Therefore by (1.0), $c = 0$ and by (2.12), (2.5) is true.

This completes Part 1. Note that $f(z)$ now satisfies condition (1.2) and therefore our theorem now follows from that of Iyer. We shall, however, give

[7] N. Levinson, loc. cit., Theorem 2. This proof does not depend on Dirichlet series gap theorems as does Bernstein's. Although the condition $\underline{\lim}_{n \to \infty} (\lambda_{n+1} - \lambda_n) > 0$ is stated in the hypothesis, the proof is unchanged with the less restrictive assumption that the index of condensation is zero.

a more direct proof of this result using (2.5). We use an interpolation formula that is also used by Iyer in his proof.

Part 2. *Here we complete the theorem.*

Let $\lambda_{-n} = -\lambda_n$ and $\mu_{-n} = -\mu_n$. Consider for any integer $N > 0$,

$$H_N(z) = \frac{f^N(z)}{A(z)\,B(z)} - \sum_{-\infty}^{\infty}{}' \frac{f^N(\lambda_n)}{A'(\lambda_n)\,B(\lambda_n)\,(z-\lambda_n)} - \sum_{-\infty}^{\infty}{}' \frac{f^N(i\mu_n)}{A(i\mu_n)\,B'(i\mu_n)\,(z-i\mu_n)}. \tag{2.13}$$

By (2.1), (2.3), and (2.5) the series converge. Clearly $A(z)B(z)H_N(z)$ is an integral function of order one. Also $A(z)B(z)H_N(z)$ vanishes at $z = \pm\lambda_n$, $z = \pm i\mu_n$, which are the zeros of $A(z)$ and $B(z)$ respectively. Thus $H_N(z)$ itself is an integral function of order one. Also by (2.5)

$$H_N(\pm re^{\pm i\alpha}) = O\left(\exp\left\{\left[N\epsilon - (a^2+b^2)^{\frac{1}{2}}\right]r\right\} + \frac{1}{r}\right).$$

Since ϵ can be taken arbitrarily small,

$$H_N(\pm re^{\pm i\alpha}) = o(1). \tag{2.14}$$

Hence $H_N(z)$ is bounded on four lines, am $z = \pm\alpha$, am $z = \pm(\pi - \alpha)$. And again by Pragmén-Lindelöf it follows that it is bounded in the entire plane, and thus a constant. But by (2.14) this constant must be zero. Thus (2.13) becomes

$$f^N(z) = A(z)\,B(z)\left(\sum_{-\infty}^{\infty}{}' \frac{f^N(\lambda_n)}{A'(\lambda_n)\,B(\lambda_n)\,(z-\lambda_n)} + \sum_{-\infty}^{\infty}{}' \frac{f^N(i\mu_n)}{B'(i\mu_n)\,A(i\mu_n)\,(z-i\mu_n)}\right). \tag{2.15}$$

By (1.0) there exists an M such that $|f(\pm\lambda_n)| < M$, $|f(\pm i\mu_n)| < M$. By (2.15) there exists a $c > 1$, independent of N such that

$$|f^N(\pm re^{\pm i\alpha})| \leqq cM^N e^{(a+b)r}.$$

Thus

$$|f(\pm re^{\pm i\alpha})| \leqq cMe^{(a+b)r/N}.$$

Since N can be chosen arbitrarily large, it follows that

$$|f(\pm re^{\pm i\alpha})| \leqq cM.$$

But this implies that $f(z)$ is bounded in the entire plane, and is therefore a constant.

3. There are many extensions of the theorem considered here. First there is the case of integral functions of order not one. Iyer obtains results for these

which may be extended to best possible results by using the same method as is used in proving our theorem.

It is also simple to remove the requirement that $\{\lambda_n\}$ and $\{\mu_n\}$ be real. It is only necessary that

$$\lim_{n\to\infty} (\mathrm{am}\, \lambda_n) = 0, \qquad \lim_{n\to\infty} (\mathrm{am}\, \mu_n) = 0.$$

There is an extension under quite general conditions to sequences not necessarily even, that is, where $\lambda_{-n} = -\lambda_n$ need not hold. Here again all changes in the proof are quite obvious.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

SOME GAP THEOREMS FOR POWER SERIES

By R. P. Boas, Jr.

1. **Introduction.** In this note we are concerned with the behavior of certain power series,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \tag{1.1}$$

on the circle of convergence (which we suppose to be the unit circle). We consider, not the convergence of the series (1.1) for $|z| = 1$, but the convergence of a suitably chosen subsequence of its partial sums. It is to be expected that such a sequence may converge under hypotheses on $f(z)$ lighter than those which ensure the convergence of the series itself; but the situation is complicated by the fact that the sequence, unlike the sequence of all the partial sums, may converge without converging to the "right" value. By way of illustration, we consider three examples. We write $s_n = \sum_{k=0}^{n} a_k$.

(i) $a_n = (-1)^n$, $f(z) = 1/(1+z)$. Then $s_{2n-1} \to 0$, but $f(z) \to \frac{1}{2}$ as $z \to 1$.

(ii) $a_0 = 1$, $a_{2m} = -a_{2m-1} = 4m + 1$ $(m = 1, 2, \cdots)$; $f(z) = 2(1+z)^{-2} - (1-z)^{-1}$. Here $s_{2n} \to 1$, but $|f(z)| \to \infty$.

(iii) $a_{3n} = a_{3n+2} = 1$, $a_{3n+1} = -2$ $(n = 0, 1, 2, \cdots)$; $f(z) = (1-z)^2/(1-z^3)$. Then $s_{3n-1} \to 0$ and $f(z) \to 0$.

Apparently the only results in the literature concerning the convergence of subsequences of partial sums are those of A. Ostrowski. Ostrowski's theorem[1] states that if there are infinitely many sufficiently long gaps in the sequence $\{a_n\}$ (if, in fact, $a_n = 0$ for $n_k < n < n_k(1+\epsilon)$, where $\epsilon > 0$ and $n_k \to \infty$), then regularity of $f(z)$ on a closed arc of $|z| = 1$ implies the uniform convergence of the sequence $s_{n_k}(z) = \sum_{n=0}^{n_k} a_n z^n$ to $f(z)$ in a domain containing that arc—in particular, then, on the arc. Our theorems resemble the result of Ostrowski in assuming the existence of gaps in the sequence $\{a_n\}$; they assume less than the regularity of $f(z)$ at points of the circle of convergence, but require restrictions on the rate of growth of the a_n. Naturally, we obtain convergence of $\{s_{n_k}(z)\}$ only for points of $|z| = 1$, not for exterior points. In one theorem (Theorem 2) we require $f(z)$ to be of bounded variation on an arc of $|z| = 1$; we obtain convergence of $\{s_{n_k}(z)\}$ on the arc under the assumption of smaller gaps than those necessary for Ostrowski's theorem. This result extends a theorem of P. Fatou,[2] which states that $a_n = o(1)$ implies convergence of (1.1) at every

Received November 17, 1937. The author is a National Research Fellow.

[1] Dienes [2], p. 358.

[2] Dienes [2], p. 467.

regular point, and the generalization given independently by P. Dienes[3] and W. H. Young,[4] in which the conclusion is convergence on every arc of $|z| = 1$ where $f(z)$ is of bounded variation. In Theorems 3 and 4, on the other hand, we require merely that $f(z)$ approaches a limit as $z \to 1$ on the real axis, but we then need gaps larger than those of Ostrowski's theorem. In Theorems 3 and 4 we actually establish equiconvergence of the sequences $\{s_{n_k}\}$ and $\{f(x_k)\}$, where $\{x_k\}$ is a particular set of points approaching unity, so that these theorems may be applied to give results of an Abelian as well as of a Tauberian character.

We derive Theorem 2 from an extension of Riemann's localization theorem for trigonometrical series; this extension seems to be of some independent interest.

2. **The localization theorem.** We shall consider a pair of functions $\varphi(n)$ and $\lambda(n)$ of the real variable n, $0 < n < \infty$; we require them to satisfy

Conditions A. $\varphi(n) > 0$; $\varphi(n) \uparrow \infty$; $\varphi(n) = O(n^q)$ $(n \to \infty)$ *for some* q, $0 < q < 1$; $\varphi(n)$ *is an integer when* n *is.*

$\lambda(n) > 0$; $\lambda(n)$ *is non-decreasing*; $\lambda(n) = O(\varphi(\frac{1}{2}n)^s)$ $(n \to \infty)$, *for some* $s > 0$.

Examples of functions satisfying Conditions A are:

$\lambda(n) = 1$, $\varphi(n)$ any function becoming infinite with arbitrary slowness and sufficient regularity;

$\lambda(n) = n^\alpha$, $\alpha > 0$; $\varphi(n) = [n^\beta]$ for any β with $0 < \beta < 1$;[5]

$\lambda(n) = \log(1 + n)$, $\varphi(n) = [\lambda(n)^\alpha]$ for any $\alpha > 0$.

We shall also consider a function $\rho(t)$ defined in relation to an interval (a, b) $(0 \leqq a < b < 2\pi)$ in a manner which we specify in what we shall call

Conditions $B_m(a, b)$. $a < \alpha < \beta < b$; $\rho(t) \equiv 1$ $(\alpha \leqq t \leqq \beta)$; $\rho(t) \equiv 0$ *on* $(0, a)$ *and* $(b, 2\pi)$; $\rho(t)$ *of class* $C^{2m+2}(0, 2\pi)$;[6] $\rho^{(n)}(\alpha) = \rho^{(n)}(\beta) = 0$ $(n = 1, 2, \cdots, 2m + 2)$; $\rho(t)$ *periodic with period* 2π.

The existence of such a function $\rho(t)$ is easily established. We write

$$D_n(t) = \frac{\sin(n + \frac{1}{2})t}{2\sin\frac{1}{2}t} \qquad (n = 0, 1, 2, \cdots) \tag{2.1}$$

($D_n(t)$ is the "Dirichlet kernel").

Given a formal trigonometrical series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad (c_0 = 0;\ c_{-n} = \bar{c}_n,\ n = 1, 2, \cdots), \tag{2.2}$$

we define its partial sums $s_k(x)$ by

$$s_k(x) = \sum_{n=-k}^{k} c_n e^{inx}.$$

[3] Dienes [1], p. 37; [2], p. 471.

[4] Young [5], p. 365.

[5] $[x]$ denotes the greatest integer $\leqq x$.

[6] $C^k(a, b)$ denotes the class of functions with continuous derivatives of order k on (a, b).

We write, formally,

$$F_m(x) = (-i)^m \sum_{n=-\infty}^{\infty}{}' c_n n^{-m} e^{inx} \qquad (m = 1, 2, \cdots);[7] \tag{2.3}$$

$F_m(x)$ is (2.2) formally integrated m times. If $c_n = O(n^{m-2})$ $(n \to \infty)$, $F_m(x)$ is actually defined by the (absolutely convergent) series (2.3).

THEOREM 1. *Let the sequence* $\{c_n\}$ *define the formal trigonometrical series* (2.2). *Let* $\varphi(n)$ *and* $\lambda(n)$ *satisfy Conditions* A, *let* $c_n = O(\lambda(n))$ $(n \to \infty)$, *and let there exist a sequence of positive integers* n_k *such that* $n_k \to \infty$, *and for* $n_k - \varphi(n_k) < n < n_k + \varphi(n_k)$, $c_n = o(1)$ $(n \to \infty)$. *If* m *is an integer greater than* s *and not less than* $qs + 2$, *if* $F_m(x)$ *is defined by* (2.3), *and if* $\rho(t)$ *satisfies Conditions* $B_m(a, b)$, *then the sequence*

$$\sum_{n=-n_k}^{n_k} c_n e^{inx} - \frac{(-1)^m}{\pi} \int_0^{2\pi} F_m(t)\rho(t) \frac{d^m}{dt^m} D_{n_k}(t - x)\, dt \tag{2.4}$$

approaches zero as $k \to \infty$, *uniformly for* $\alpha \leqq x \leqq \beta$.

This theorem may be interpreted as stating that the behavior of the sequence $\{s_{n_k}(x)\}$ at a point or in an interval depends only on the behavior of $F_m(x)$ in a neighborhood of that point or interval.

The proof of Theorem 1 is modelled on the proof of Riemann's localization theorem. We follow the simplified presentation given by L. Neder.[8]

We require two lemmas.

LEMMA 1. *If* $\sigma(t)$ *is periodic of period* 2π, *and of class* $C^m(-\infty, \infty)$, *the Fourier coefficients* σ_n *of* $\sigma(t)$ *satisfy*

$$|\sigma_n| \leqq A\, |n|^{-m} \qquad (n = \pm 1, \pm 2, \cdots), \tag{2.5}$$

where $A = \max_{0 \leqq t \leqq 2\pi} |\sigma^{(m)}(t)|$.

We have

$$\sigma_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-int} \sigma(t)\, dt. \tag{2.6}$$

Integrating by parts m times, we find

$$\sigma_n = \frac{(-i)^m}{2\pi n^m} \int_0^{2\pi} e^{-int} \sigma^{(m)}(t)\, dt, \qquad |\sigma_n| \leqq A\, |n|^{-m}.$$

LEMMA 2. *Let the hypotheses of Theorem* 1 *be satisfied. Let* $\sigma(t)$ *be an integrable function of period* 2π, *having Fourier coefficients* σ_n *defined by* (2.6) *and satisfying*

$$\begin{aligned} &|\sigma_n| \leqq A\, |n|^{-m-1} \qquad (n = \pm 1, \pm 2, \cdots), \\ &|\sigma_0| \leqq A. \end{aligned} \tag{2.7}$$

[7] The accent on the summation sign indicates the omission of the term for which $n = 0$.
[8] Neder [4], pp. 117–124.

Then

$$J_{n_k} = n_k^m \int_0^{2\pi} F_m(x+t)\sigma(t) \begin{matrix}\cos\\ \sin\end{matrix}\, n_k t\, dt \to 0 \qquad (k \to \infty),$$

uniformly for all x. If $\sigma(t)$ depends on x, the conclusion still holds on any set of points for which A can be chosen independent of x.

Because $m \geqq qs + 2$, $F_m(x)$ is defined. To simplify the notation, we write N for n_k. We have, with $\alpha = 0$ or $-\pi/2$,

$$J_N = \tfrac{1}{2}N^m \int_0^{2\pi} F_m(x+t)\sigma(t)(e^{iNt+i\alpha} + e^{-iNt-i\alpha})\, dt.$$

We replace $F_m(x+t)$ by its defining series,

$$F_m(x+t) = (-i)^m \sum_{k=-\infty}^{\infty}{}' c_k k^{-m} e^{ikz},$$

integrate term by term, and obtain, using (2.7),[9]

$$J_N = \tfrac{1}{2}(-iN)^m \sum_{k=-\infty}^{\infty}{}' c_k k^{-m} \left\{ e^{ikx+i\alpha} \int_0^{2\pi} e^{i(k+N)t}\sigma(t)\, dt + e^{ikx-i\alpha} \int_0^{2\pi} e^{i(k-N)t}\sigma(t)\, dt \right\},$$

$$\begin{aligned}
|J_N| &\leqq \tfrac{1}{2}|N|^m \sum_{k=-\infty}^{\infty}{}' |c_k|\,|k|^{-m}(|\sigma_{k+N}| + |\sigma_{k-N}|) \\
&\leqq AN^m \sum_{k=1}^{\infty}{}'' |c_k|\, k^{-m}\{(N+k)^{-m-1} + |N-k|^{-m-1}\} + 2A|c_N| \\
&\leqq 2AN^m \sum_{k=1}^{\infty}{}'' |c_k|\, k^{-m}|N-k|^{-m-1} + 2A|c_N| \\
&= 2AN^m \left(\sum_{1}^{[N/2]} + \sum_{[N/2]+1}^{N-\varphi(N)} + \sum_{N-\varphi(N)+1}^{N-1} + \sum_{N+1}^{N+\varphi(N)-1} + \sum_{N+\varphi(N)}^{\infty} \right) + 2A|c_N| \\
&= S_1 + S_2 + S_3 + S_4 + S_5 + 2A|c_N|.
\end{aligned}$$

Since $\varphi(N) = o(N)$, we may (and shall) suppose that $N - \varphi(N) > [N/2]$. We have, for some constants $B > 0$, $C > 0$, $|c_n| < B\lambda(n) < Cn^{qs}$ $(n = 1, 2, \cdots)$; and, for $N - \varphi(N) < n < N + \varphi(N)$, $|c_n| < \epsilon(n)$, where $\lim_{n\to\infty} \epsilon(n) = 0$.

Since $m - qs \geqq 2$, we have

$$\begin{aligned}
S_1 &\leqq 2ACN^m \sum_{k=1}^{[N/2]} k^{qs-m}(N-k)^{-m-1} \\
&\leqq 2ACN^m(2/N)^{m+1} \sum_{k=1}^{\infty} k^{-2} \\
&= o(1) \qquad\qquad (N \to \infty).
\end{aligned}$$

[9] The double accent on the summation sign below indicates the omission of the term for which $k = N$.

Since $\lambda(k)$ is non-decreasing, and $m > s$,

$$S_2 \leqq 2^{m+1}AB \sum_{k=[N/2]+1}^{N-\varphi(N)} \lambda(k)(N-k)^{-m-1}$$

$$\leqq 2^{m+1}AB\lambda(N) \sum_{n=\varphi(N)}^{\infty} n^{-m-1}$$

$$\leqq 2^{m+1}AB\lambda(N)\{\varphi(N)-1\}^{-m}$$

$$= o(1) \qquad (N \to \infty).$$

Since $N = O(N - \varphi(N))$ $(N \to \infty)$,

$$S_3 \leqq 2AN^m \sum_{k=N-\varphi(N)+1}^{N-1} \epsilon(k)k^{-m}(N-k)^{-m-1}$$

$$\leqq 2A\{N/(N-\varphi(N)+1)\}^m \sum_{n=1}^{\varphi(N)-1} \epsilon(N-n)n^{-m-1}$$

$$= o(1) \qquad (N \to \infty).$$

Next we have

$$S_4 \leqq 2A \sum_{k=N+1}^{N+\varphi(N)-1} \epsilon(k)(k-N)^{-m-1}$$

$$= 2A \sum_{n=1}^{\varphi(N)-1} \epsilon(n+N)n^{-m-1}$$

$$= o(1) \qquad (N \to \infty).$$

Finally,

$$S_5 \leqq 2N^m AB \sum_{k=N+\varphi(N)}^{\infty} \lambda(k)k^{-m}(k-N)^{-m-1}$$

$$\leqq 2N^m AB\{N+\varphi(N)\}^{-m} \sum_{k=\varphi(N)}^{\infty} k^{-m-1}\lambda(k+N)$$

$$\leqq 2AB \sum_{k=\varphi(N)}^{\infty} k^{-m-1}\lambda(k+N).$$

But $\varphi(N)$ is non-decreasing, and hence for $k \geqq \varphi(N)$, $\varphi^{-1}(k) \geqq N$. Since $\lambda(n)$ is non-decreasing, we have

$$S_5 \leqq 2AB \sum_{k=\varphi(N)}^{\infty} k^{-m-1}\lambda(k+\varphi^{-1}(k)). \tag{2.8}$$

For sufficiently large k, since $q < 1$,

$$k > \varphi(k),$$

$$\varphi^{-1}(k) > k,$$

$$k + \varphi^{-1}(k) < 2\varphi^{-1}(k),$$

$$\lambda(k + \varphi^{-1}(k)) \leqq \lambda(2\varphi^{-1}(k)) = O\{\varphi(\varphi^{-1}(k))^s\} = O(k^s).$$

Relation (2.8) now gives, for some $D > 0$,

$$S_5 \leqq 2ABD \sum_{k=\varphi(N)}^{\infty} k^{-m+s-1} = o(1) \qquad (N \to \infty),$$

since $m > s$. This completes the proof of Lemma 2.

We now prove Theorem 1. A simple computation shows that

$$s_{n_k}(x) = \frac{(-1)^m}{\pi} \int_0^{2\pi} F_m(x + t) \frac{d^m}{dt^m} D_{n_k}(t)\, dt.$$

The difference (2.4) may then be written

$$\frac{(-1)^m}{\pi} \int_0^{2\pi} F_m(x + t)(1 - \rho(x + t)) \frac{d^m}{dt^m} D_{n_k}(t)\, dt. \tag{2.9}$$

Now, if we set $\Delta_n(t) = \frac{1}{2} \sin nt \operatorname{ctn} \frac{1}{2}t$, we have $D_n(t) = \Delta_n(t) + \frac{1}{2} \cos nt$. Hence we may write (2.9) as the sum of

$$\frac{(-1)^m}{\pi} \int_0^{2\pi} F_m(x + t)(1 - \rho(x + t)) \frac{d^m}{dt^m} \Delta_{n_k}(t)\, dt \tag{2.10}$$

and

$$\frac{(-1)^m}{2\pi} \int_0^{2\pi} F_m(x + t)(1 - \rho(x + t)) \frac{d^m}{dt^m} \cos n_k t\, dt. \tag{2.11}$$

We set $\omega(t) = \frac{1}{2} \operatorname{ctn} \frac{1}{2}t$; $\omega(t)$ is of class C^∞ for $t \not\equiv 0 \pmod{2\pi}$, and we have

$$\Delta_{n_k}(t) = \omega(t) \sin n_k t,$$

$$\begin{aligned} \frac{d^m}{dt^m} \Delta_{n_k}(t) &= \sum_{j=0}^{m} \binom{m}{j} \omega^{(m-j)}(t)(\sin n_k t)^{(j)} \\ &= \sum_{j=0}^{m} A_j n_k^j \omega^{(m-j)}(t) \begin{matrix}\cos \\ \sin\end{matrix}\, n_k t, \end{aligned} \tag{2.12}$$

where the A_j are constants independent of n_k.

We define functions $\sigma(t)$ (depending on x and p) by

$$\sigma(t) = (1 - \rho(x + t))\, \omega^{(p)}(t) \qquad (t \neq 0; 0 \leqq p \leqq m),$$

$$\sigma(0) = 0.$$

Since $\rho(t)$ satisfies Conditions $B_m(a, b)$, each $\sigma(t)$ is of class C^{m+1} for $\alpha \leqq x \leqq \beta$; and $|\sigma^{(m+1)}(t)|$ has a finite upper bound independent of x, $\alpha \leqq x \leqq \beta$. By Lemma 1, then, the Fourier coefficients of $\sigma(t)$ satisfy (2.7), with A independent of x, $\alpha \leqq x \leqq \beta$. By Lemma 2,

$$J_{n_k} = n_k^m \int_0^{2\pi} F_m(x + t)\sigma(t) \frac{\cos}{\sin} n_k t\, dt \to 0 \qquad (k \to \infty)$$

uniformly for $\alpha \leqq x \leqq \beta$; by (2.12), the expression (2.10) has the same property; and (2.11) is a constant multiple of J_{n_k} (with $p = 0$). Hence (2.9), which is the same as (2.4), also approaches zero as $k \to \infty$, uniformly for $\alpha \leqq x \leqq \beta$.

3. Application to power series.

THEOREM 2. *Let* $f(z) = \sum_{n=0}^{\infty} a_n z^n$, $\overline{\lim}_{n\to\infty} |a_n|^{1/n} = 1$. *Let* $\varphi(n)$ *and* $\lambda(n)$ *satisfy Conditions* A; *let* $a_n = O(\lambda(n))$ $(n \to \infty)$; *and let there exist a sequence of integers* n_k *such that* $n_k \to \infty$, *and for* $n_k - \varphi(n_k) < n < n_k + \varphi(n_k)$, $a_n = o(1)$ $(n \to \infty)$. *If, for* $a < \theta < b$, $\lim_{r\to 1} f(re^{i\theta}) = g(\theta)$, *and* $g(\theta)$ *is a normalized*[10] *function, of bounded variation on every* (α, β) $(a < \alpha < \beta < b)$, *then* $s_{n_k}(\theta) = \sum_{n=0}^{n_k} a_n e^{in\theta}$ *converges to* $g(\theta)$ $(k \to \infty)$; *if* $f(re^{i\theta})$ *converges uniformly for* $\alpha \leqq \theta \leqq \beta$, *so does* $s_{n_k}(\theta)$.

In particular, then, if $f(re^{i\theta})$ is regular for $r = 1$, $a < \theta < b$, $s_{n_k}(\theta) \to f(e^{i\theta})$, uniformly for $\alpha \leqq \theta \leqq \beta$.

The scope of Theorem 2 will perhaps be clearer if we give some examples of how the hypotheses of the theorem can be satisfied. Let us call $\varphi(n_k)$ "the length of the gap at n_k". If

(i) $a_n = O(1)$ and the length of the gap at n_k becomes infinite; or

(ii) $a_n = O(n^\alpha)$ for some $\alpha > 0$ (however large), and the length of the gap at n_k increases as rapidly as n_k^β, for some $\beta > 0$ (however small); or

(iii) $a_n = O(\log n)$, and the length of the gap at n_k increases as rapidly as some positive power of $\log n_k$;

then $s_{n_k}(\theta)$ approaches $f(e^{i\theta})$ on any arc of $|z| = 1$ where $f(z)$ is of bounded variation.

COROLLARY. *Let* $f(z) = \sum_{k=0}^{\infty} a_k z^{\nu_k}$, $\overline{\lim}_{k\to\infty} |a_k|^{1/\nu_k} = 1$. *Let* $\varphi(n)$ *and* $\lambda(n)$ *satisfy Conditions* A; *let* $a_k = O(\lambda(\nu_k))$ $(k \to \infty)$; *let* $\varphi(\nu_k) = O(\nu_{k+1} - \nu_k)$ $(k \to \infty)$. *If* $\overline{\lim}_{k\to\infty} |a_k| \neq 0$, $f(z)$ *cannot be of bounded variation on any arc of* $|z| = 1$.[11]

With the hypotheses of the corollary, the Fabry gap theorem[12] gives $|z| = 1$ as a natural boundary (even without any restriction on the order of magnitude of the a_k). With the order of magnitude of the coefficients restricted, we obtain a more precise result.

[10] $g(\theta)$ is normalized if $g(\theta) = \frac{1}{2}[g(\theta+) + g(\theta-)]$, $a < \theta < b$.

[11] To apply Theorem 2 to $f(z)$, we take $n_k = [\frac{1}{2}(\nu_k + \nu_{k+1})]$, and (if necessary) replace $\varphi(n)$ by a constant multiple of $\varphi(n)$.

[12] Dienes [2], p. 376.

To prove Theorem 2, we let

$$f(r, \theta) = \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta} \qquad (c_{-n} = \bar{c}_n) \tag{3.1}$$

be the real or imaginary part of $f(re^{i\theta})$. There is no loss of generality from assuming that $c_0 = 0$, and that $0 \leqq a < b < 2\pi$. Let $\lim_{r\to 1} f(r, \theta) = h(\theta)$ on (a, b), and let $h(\theta)$ be extended outside (a, b) so that it is integrable on $(0, 2\pi)$ and periodic with period 2π, with $\int_0^{2\pi} h(\theta)\, d\theta = 0$. Let the interval (α, β) be given, and choose a', $a < a' < \alpha$.

Since the series in (3.1) is uniformly convergent for any r, $0 < r < 1$, we may integrate it term by term, m times, where $m \geqq qs + 2$ and $m > s$, obtaining for any θ

$$(-i)^m \sum_{n=-\infty}^{\infty}{}' c_n n^{-m} r^{|n|} e^{in\theta} + P_{m-1}(r, \theta) = \frac{1}{(m-1)!} \int_{a'}^{\theta} (\theta - u)^{m-1} f(r, u)\, du, \tag{3.2}$$

where $P_{m-1}(r, \theta)$ is for each r a polynomial in θ of degree $\leqq m - 1$.

The series in (3.2) is the real or imaginary part of the series

$$\sum_{n=1}^{\infty} a_n n^{-m} z^n; \tag{3.3}$$

since, by our hypotheses on the a_n, the series $\sum_{n=1}^{\infty} | a_n | n^{-m}$ converges, it follows by Abel's theorem that the function (3.3) approaches a limit as $| z | \to 1$, uniformly for all θ; its real and imaginary parts have the same property. The series in (3.2) therefore approaches

$$(-i)^m \sum_{n=-\infty}^{\infty}{}' c_n n^{-m} e^{in\theta}, \tag{3.4}$$

uniformly for all θ.

For $a' \leqq \theta \leqq b'$ $(b > b' > \beta)$, $g(\theta)$ is a function of bounded variation and hence bounded. For $0 \leqq r < 1$, $f(re^{ia'})$ and $f(re^{ib'})$ are bounded, uniformly with respect to r. Hence $f(re^{i\theta})$ is bounded in the sector $0 \leqq r \leqq 1$, $a' \leqq \theta \leqq b'$, its real and imaginary parts have the same property, and

$$\lim_{r\to 1} \int_{a'}^{\theta} f(r, u)\, du = \int_{a'}^{\theta} h(u)\, du,$$

boundedly for $a' \leqq \theta \leqq b'$. Integrating $m - 1$ times more, we find that

$$\lim_{r\to 1} \frac{1}{(m-1)!} \int_{a'}^{\theta} (\theta - u)^{m-1} f(r, u)\, du = \frac{1}{(m-1)!} \int_{a'}^{\theta} (\theta - u)^{m-1} h(u)\, du,$$

for $a' \leqq \theta \leqq b'$. It now follows from (3.2) that $\lim_{r\to 1} P_{m-1}(r, \theta)$ exists for $a' \leqq$

$\theta \leqq b'$ and is therefore a polynomial $P_{m-1}(\theta)$ for $a' \leqq \theta \leqq b'$. We then have, for $a' \leqq \theta \leqq b'$,

$$F_m(\theta) \equiv (-i)^m \sum_{n=-\infty}^{\infty}{}' c_n n^{-m} e^{in\theta} = -P_{m-1}(\theta) + \frac{1}{(m-1)!}\int_{a'}^{\theta} (\theta - u)^{m-1} h(u)\, du.$$

Let the Fourier series of $h(\theta)$ be

$$\sum_{n=-\infty}^{\infty} \gamma_n e^{in\theta} \qquad (\gamma_0 = 0); \tag{3.5}$$

let

$$\Phi_m(\theta) = (-i)^m \sum_{n=-\infty}^{\infty}{}' \gamma_n n^{-m} e^{in\theta};$$

since term by term integration of (3.5) is legitimate,

$$\Phi_m(\theta) = Q_{m-1}(\theta) + \frac{1}{(m-1)!}\int_{a'}^{\theta} (\theta - u)^{m-1} h(u)\, du \qquad (a' \leqq \theta \leqq b'),$$

where $Q_{m-1}(\theta)$ is a polynomial of degree $\leqq m - 1$. Therefore

$$G_m(\theta) \equiv F_m(\theta) - \Phi_m(\theta) = R_{m-1}(\theta) \qquad (a' \leqq \theta \leqq b'),$$

where $R_{m-1}(\theta)$ is a polynomial of degree $\leqq m - 1$.

The trigonometrical series

$$\sum_{n=-\infty}^{\infty} (c_n - \gamma_n) e^{in\theta}$$

satisfies the hypotheses of Theorem 1. If $\rho(t)$ satisfies Conditions $B_m(a', b')$, we shall have, by Theorem 1,

$$\lim_{k\to\infty} \left\{ \sum_{n=-n_k}^{n_k} (c_n - \gamma_n) e^{in\theta} - \frac{(-1)^m}{\pi} \int_0^{2\pi} G_m(t)\rho(t) \frac{d^m}{dt^m} D_{n_k}(t - \theta)\, dt \right\} = 0 \tag{3.6}$$

uniformly for $\alpha \leqq \theta \leqq \beta$. We shall show that the limit of the integral in (3.6) is zero.

We suppose, as we may, that $\rho^{(n)}(a') = \rho^{(n)}(b') = 0$ $(n = 0, 1, \cdots, m - 1)$; then $G_m(t)\rho(t)$ is of class $C^m(a', b')$, and its derivatives of order $\leqq m - 1$ vanish at a' and at b'. We integrate the integral in (3.6) m times by parts, obtaining

$$\int_{a'}^{b'} D_{n_k}(t - \theta) \frac{d^m}{dt^m} \{G_m(t)\rho(t)\}\, dt. \tag{3.7}$$

Since $\{G_m(t)\rho(t)\}^{(m)}$ is of class $C'(a', b')$, the Dirichlet integral (3.7) converges on $a' \leqq \theta \leqq b'$ and uniformly on (α, β) to $\{G_m(t)\rho(t)\}^{(m)}$; on (α, β), $\rho(t) \equiv 1$ and $\{G_m(t)\}^{(m)} = 0$, so that this limit is zero.

Referring to (3.6), and using the fact that (3.5) converges to $h(\theta)$, $a' < \theta < b'$, we see that

$$\lim_{k\to\infty} \sum_{n=-n_k}^{n_k} c_n e^{in\theta} = \lim_{k\to\infty} \sum_{n=-n_k}^{n_k} \gamma_n e^{in\theta} = h(\theta) \qquad (\alpha \leqq \theta \leqq \beta). \tag{3.8}$$

The relation (3.8) holds uniformly if $f(re^{i\theta}) \to g(\theta)$ uniformly, since $h(\theta)$ is then continuous as well as of bounded variation. Since these results are true whether $f(r, \theta)$ is the real or imaginary part of $f(re^{i\theta})$, we obtain our conclusion.

4. **Theorems with large gaps.** In §3 we considered power series with strong restrictions on their behavior on the circle of convergence, and small gaps in the coefficients. We now consider power series subjected to weak restrictions on the circle of convergence, but with large gaps.

Let $\lambda(n)$ be a positive function of positive integers n, with $\overline{\lim}_{n\to\infty} \lambda(n)^{1/n} \leqq 1$. We shall consider the class of power series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{4.1}$$

with $0 \leqq x < 1$, subject to

$$\overline{\lim_{n\to\infty}} \, | a_n |^{1/n} = 1,$$

and

$$a_n = o(\lambda(n)) \qquad (n \to \infty). \tag{4.2}$$

We shall call this class $C(\lambda(n))$.

We write, given a function (4.1), $s_k = \sum_{n=0}^{k} a_n$.

THEOREM 3. *For every class $C(\lambda(n))$ there exists a positive function $\varphi(n)$ such that if $f(x)$ belongs to $C(\lambda(n))$ and if there is a sequence $\{n_k\}$ with $\lim_{k\to\infty} n_k = \infty$ and with $a_n = 0$ for $n_k < n < n_k + \varphi(n_k)$, then for a suitable sequence $\{x_n\}$, with $\lim_{n\to\infty} x_n = 1$,*

$$\lim_{k\to\infty} \{f(x_k) - s_{n_k}\} = 0. \tag{4.3}$$

Thus, if $f(x_k)$ approaches a limit s $(k \to \infty)$ (in particular, if $f(x) \to s$ as $x \to 1$), $s_{n_k} \to s$; if $s_{n_k} \to s$, $f(x_k) \to s$ (but $f(x)$ does not necessarily approach s, as is easily shown by examples).

THEOREM 4. *For the class $C(n^{r-2})$ $(r > 1)$, we may take $x_k = 1 - n_k^{-r}$ and $\varphi(n) = [r^2 n^r \log n]$.*

If $r \leqq 1$, $C(n^{r-2})$ becomes the class of series with $a_n = o(n^{-1})$, for which Tauber's theorem[13] gives a better result than Theorem 4 would.

If $n_k + \varphi(n_k) = n_{k+1}$, so that (4.1) has the form $\sum_{k=0}^{\infty} a_{n_k} x^{n_k}$, Theorem 4 fails to contain the "high indices theorem" of Hardy and Littlewood;[14] however, it gives, for the series to which it applies, more information than the high indices theorem does.

[13] Dienes [2], p. 465.
[14] Ingham [3].

Theorem 4 could be sharpened somewhat at the expense of additional complication in statement, but since it is by no means certain that even better results could not be obtained by more powerful methods, we leave the statement in the simpler form.

To establish Theorem 3, we write

$$f(x_k) - s_{n_k} = \sum_{n=n_k+\varphi(n_k)}^{\infty} a_n x_k^n - \sum_{n=0}^{n_k} a_n(1 - x_k^n) = S_1 - S_2. \tag{4.4}$$

Then

$$|S_2| \leqq (1 - x_k) \sum_{n=0}^{n_k} n\,|a_n| = o\left\{(1 - x_k) \sum_{n=1}^{n_k} n\lambda(n)\right\} \qquad (k \to \infty). \tag{4.5}$$

Let $\omega(n)$ be a positive function such that $\omega(n) \geqq 1$, $\omega(n) \to \infty$ as $n \to \infty$, and

$$\omega(n) \geqq \sum_{p=0}^{n} p\lambda(p) \qquad (n = 1, 2, \cdots).$$

Then, by (4.5),

$$S_2 = o\{(1 - x_k)\omega(n_k)\} \qquad (k \to \infty);$$

$$S_2 = o(1)$$

if $x_k = 1 - 1/\omega(n_k)$.

We have

$$S_1 = \sum_{n=n_k+\varphi(n_k)}^{\infty} n^{-1} a_n n x_k^n.$$

Let k be so large that $|a_n| < \epsilon\lambda(n)$, $n > n_k$, where $\epsilon > 0$ is arbitrary. Then

$$|S_1| \leqq \frac{\epsilon}{n_k + \varphi(n_k)} \sum_{n=n_k+\varphi(n_k)}^{\infty} n\lambda(n) x_k^n$$

$$\leqq \frac{\epsilon\psi(x_k)}{n_k + \varphi(n_k)},$$

where

$$\psi(x) = \sum_{n=1}^{\infty} n\lambda(n) x^n.$$

We have $x_k = 1 - 1/\omega(n_k)$; if then

$$\varphi(n) = \psi(1 - 1/\omega(n)),$$

we have $|S_1| \leqq \epsilon$. This proves Theorem 3.

To establish Theorem 4, we start again from (4.4); when $\lambda(n) = n^{r-2}$, we may take $\omega(n) = n^r$; and if $x_k = 1 - n_k^{-r}$, $S_2 = o(1)$ $(k \to \infty)$. We now estimate S_1

more precisely than before. Take k so that for $n > n_k$, $| a_n | < n^{r-2}$, and write $N = n_k$, $x = x_k$. Then

$$(4.6)\quad \begin{aligned} | S_1 | &\leqq x^{N+\varphi(N)-1} \sum_{n=N+\varphi(N)}^{\infty} n^{r-2} x^{n-N-\varphi(N)+1} \\ &\leqq x^{\varphi(N)} \sum_{n=0}^{\infty} x^n \{n + N + \varphi(N)\}^{r-2}. \end{aligned}$$

Suppose first that $r \geqq 2$. Then since $N \leqq \varphi(N)$, we have

$$| S_1 | \leqq x^{\varphi(N)} \left(\sum_{n=0}^{2\varphi(N)} + \sum_{n=2\varphi(N)}^{\infty} \right) x^n \{n + 2\varphi(N)\}^{r-2} = S_1' + S_1''.$$

Then

$$S_1' \leqq 4^{r-2} \varphi(N)^{r-2} x^{\varphi(N)} (1 - x)^{-1},$$

$$S_1'' \leqq 2^{r-2} x^{\varphi(N)} \sum_{n=1}^{\infty} x^n n^{r-2};$$

since

$$\sum_{n=0}^{\infty} \binom{n + r - 2}{n} x^n = (1 - x)^{-r+1}$$

and

$$n^{r-2} \sim \Gamma(r - 1) \binom{n + r - 2}{n} \qquad (n \to \infty),$$

we have

$$S_1'' \leqq A x^{\varphi(N)} (1 - x)^{-r+1},$$

where A is a suitable constant. Writing $\varphi(N) = N^r \mu(N)$, and remembering that $1 - x = N^{-r}$, we have (with a new A)

$$| S_1 | \leqq A N^{r^2 - r} (1 - N^{-r})^{N^r \mu(N)} \{1 + \mu(N)^{r-2}\}.$$

Under our hypotheses, $\mu(N) \sim r^2 \log N$ $(N \to \infty)$; an elementary computation shows that

$$(1 - N^{-r})^{N^r \mu(N)} N^{r^2 - r} \mu(N)^{r-2} = o(1) \qquad (N \to \infty);$$

thus $S_1 = o(1)$ $(k \to \infty)$ if $r \geqq 2$.

If $1 < r < 2$, we have from (4.6)

$$\begin{aligned} | S_1 | &\leqq x^{\varphi(N)} \sum_{n=0}^{\infty} x^n (n + 2N)^{r-2} \\ &\leqq 2^{r-2} N^{r-2} x^{\varphi(N)} (1 - x)^{-1} \\ &= 2^{r-2} N^{2r-2} (1 - N^{-r})^{N^r \mu(N)}. \end{aligned}$$

It is easily verified that the last expression is $o(1)$ if $\mu(N) \geqq (2r - 1) \log N$ for sufficiently large N; but $\mu(N) \sim r^2 \log N > (2r - 1) \log N$ (since $1 < r < 2$). This completes the proof.

References

1. P. Dienes, *Leçons sur les singularités des fonctions analytiques*, Paris, 1913.
2. P. Dienes, *The Taylor Series*, Oxford, 1931.
3. A. E. Ingham, *On the "high-indices theorem" of Hardy and Littlewood*, Quarterly Journal of Mathematics (Oxford series), vol. 8 (1937), pp. 1–7.
4. L. Neder, *Zur Theorie der trigonometrischen Reihen*, Mathematische Annalen, vol. 84 (1921), pp. 117–136.
5. W. H. Young, *On restricted Fourier series and the convergence of power series*, Proceedings of the London Mathematical Society, (2), vol. 17 (1916–18), pp. 353–366.

Princeton University.

ON ABSOLUTELY CONTINUOUS TRANSFORMATIONS IN THE PLANE

By Tibor Radó

1. Introduction

1.1. We shall be concerned with continuous transformations T given in the form

$$T: x = x(u, v), \qquad y = y(u, v),$$

where $x(u, v)$, $y(u, v)$ are continuous in the closed *fundamental square* $S_0: 0 \leqq u \leqq 1, 0 \leqq v \leqq 1$. *We do not assume that T is bi-unique.* Let us use K to denote any finite system of closed squares s in S_0, without common interior points. Let $\bar{s}$ denote the image of s and $|\bar{s}|$ the measure[1] of $\bar{s}$. If there exists a finite constant M such that

$$\sum_{s \in K} |\bar{s}| < M$$

for every system K, then T is of *bounded variation in the sense of Banach.* If for every $\epsilon > 0$ there exists an $\eta = \eta(\epsilon) > 0$ such that

$$\sum_{s \in K} |\bar{s}| < \epsilon$$

for every system K which satisfies the condition

$$\sum_{s \in K} |s| < \eta(\epsilon),$$

then T is *absolutely continuous in the sense of Banach* (Banach [2]).

1.2. A real function $f(u)$ of the real variable u, continuous in an interval $a \leqq u \leqq b$, gives rise to a one-dimensional continuous transformation $x = f(u)$. The theory of the transformations T, defined in §1.1, appears thus as a two-dimensional generalization of the theory of functions of a single variable which are of bounded variation and absolutely continuous, respectively. In the one-dimensional case, we know that if $f(u)$ is of bounded variation, then $f'(u)$ exists almost everywhere in the interval $a \leqq u \leqq b$ and is summable there. Furthermore, if $f(u)$ is absolutely continuous, then we have the fundamental identities

$$\int_a^b f'(u)du = f(b) - f(a), \tag{1}$$

$$\int_a^b |f'(u)| du = V(f; a, b), \tag{2}$$

Received November 29, 1937.

[1] If E is a measurable set, then $|E|$ denotes the measure of E.

where $V(f; a, b)$ denotes the total variation of $f(u)$ in $a \leqq u \leqq b$. Banach himself and subsequently Schauder (Banach [2], Schauder [1]) obtained a number of important results concerned with the two-dimensional theorems which correspond to the one-dimensional theorems stated above. In the two-dimensional case, $f'(u)$ and $|f'(u)|$ are replaced by certain *generalized Jacobians*, while the right sides of (1) and (2) are replaced by what may be called *the signed area* and *the absolute area* of the image of the fundamental square S_0, these areas being defined as the integrals of certain multiplicity functions. One of the main results of Schauder was extended in an earlier paper of the author (Radó [2]). The present paper contains a number of further contributions to the theory.

1.3. We now make a few general statements concerning its contents. The methods of Banach and Schauder are based essentially on geometrical ideas, and one of the main purposes of the present paper is to explore more fully the geometrical background of the theory. Our methods will depend upon the following geometrical tools in addition to those developed and used by Banach and Schauder. We shall take advantage of certain simple facts concerning the topological index (see Part 3) which were not considered in their work.[2] Simultaneously with the image of the fundamental square S_0 under the transformation T we shall consider *the kernel* of this image (see Part 6). The kernel represents, in a sense, *the stable portion of the image*. The idea of the kernel has been used already in an earlier paper of the author (Radó [1]) as a possibly useful substitute for certain very complicated geometrical conceptions developed by Geöcze in his work on the area of continuous surfaces.[3] Finally, we shall utilize, in Part 10, the notion of *total differentiability* in a manner suggested by the investigations of Stepanoff on *approximate differentiability* (cf. Saks [1], Chapter 9).

1.4. While the geometrical facts used by Banach and Schauder remain valid in n-dimensional Euclidean space, this is not always the case with those used in this paper, and the question of the validity of some of our results in spaces of higher dimension may lead to interesting problems. But in the two-dimensional case our methods yield new information concerning a variety of fundamental questions and lead to a more complete theory, both in the way of more precise answers to problems partially solved previously and in the way of new problems which become accessible to discussion. On account of the number of definitions and topics it seems impractical to give at this time a detailed description of the contents of this paper, and the reader is referred to §§3.2, 5.3, 5.5, 5.6, 6.2, 6.8, 6.11, 7.8, 7.10, 7.11, 7.14, 8.6–8.9, 9.2, 10.13 for quick informa-

[2] See Kerékjártó [1] for the topological facts used in the sequel.

[3] See Radó [1] for references to the work of Geöcze.

tion concerning individual items which may be considered as new contributions to the Banach-Schauder theory.

1.5. Two-dimensional generalizations of the identities (1), (2) were investigated by many authors. While the principal formulas of the various theories are very similar in appearance, they may differ considerably in their meaning according to the definitions adopted. All the quantities appearing in our theorems have a definite geometrical meaning, but this is not generally the case in other theories, and therefore a comparison of results would be possible only on the basis of an adequate discussion of the geometrical background of the various theories. Such a discussion may lead to interesting and difficult problems, particularly in connection with the work of Young [1], [2], Burkill [1], [2], and Caccioppoli [1]. We shall restrict ourselves, in the body of the paper, to references to previous results which are directly comparable to ours.

1.6. While our approach throughout the paper will be essentially geometrical, we shall make extensive use of the modern theory of differentiation and integration.[4] As might be expected, it will happen at times that we shall find it convenient or necessary to modify slightly certain well-known results and methods of the general theory. This will be the case in Part 2, in connection with the theory of non-additive set-functions due to Banach, and in Part 10 in connection with the work of Stepanoff on approximate differentiability. Our more or less detailed presentation of such modifications is not meant to imply any claims to originality, and we give some details only to keep within reasonable bounds the amount of work left to the reader.

1.7. In view of the number of permanent notations which will be used, we give here a list of notations which the reader is requested to consult whenever in doubt about the meaning of a symbol. In the following list each symbol is given together with a reference to the section where its meaning is explained.

Point-sets: K, §1.1; S_0, s, §2.2; O, C, §2.7; $\mathfrak{S}_\infty$, B_0, $\bar{B}_0$, §4.1; $\mathfrak{E}_s$, §4.11; I, $\bar{I}$, I_k, $\bar{I}_k$, I^*, $\bar{I}^*$, b, $\bar{b}$, §5.1; $\mathfrak{R}_k$, §6.2.

Functions and functionals: $N_E(z)$, $N(z)$, §4.1; $g_s(z)$, §4.4; $i(w)$, $n_s(z)$, $n(z)$, §5.1; $\psi_j(z)$, $\gamma_s(z)$, §5.6; $\kappa(z)$, §6.8; $G(s)$, $D(w)$, §7.1; $J(w)$, §7.10; $\Gamma(s)$, $\Gamma'(w)$, §8.9; $\xi(u, v; u_0, v_0)$, $\eta(u, v; u_0, v_0)$, §10.2.

Other symbols: D_j, $K_j(O)$, $K_j(C)$, §2.7; $D(s)$, §2.10; T, §4.1; p_i, §4.3; $\| T_1, T_2 \|$, §6.1; BV, AC, §7.1; $C(\{s_j\}; u_0, v_0)$, §10.4; $C(T; u_0, v_0)$, §10.6; $C(s; u_0, v_0; E)$, §10.10.

1.8. To simplify the formulas, we shall omit the symbols $dxdy$ and $dudv$ in writing double integrals. Functions in the xy-plane will be assumed to be defined in the whole plane and equal to zero outside of a sufficiently large circle.

[4] See Saks [1] for information concerning the theory of functions of real variables.

The range of integration will be indicated only in case it is different from the whole plane. The Lebesgue integral is used throughout the paper.

2. On functions of squares

2.1. We shall state in this section certain results of Banach [1] concerned with real-valued *functions of sets* $F(E)$, in a form convenient for our purposes. Some of the functions $F(E)$ which arise in geometrical applications are defined originally only for certain very simple figures (like a region bounded by a simple closed polygon), and the range of definition of $F(E)$ fails then to possess the closure properties required in the general theory developed by Banach[5] (see Banach [1]). In such cases, it may be possible to extend the range of definition of $F(E)$, but one may desire to avoid the labor involved in the process of extension and the loss of simplicity in the geometrical interpretation. It may also happen, as in the present paper, that one needs only facts which depend solely upon the values of $F(E)$ on *squares*. It might therefore be of interest to observe that the definitions and results of Banach, as far as they are needed in the sequel, can be modified in such a way as to involve only the values of $F(E)$ on squares, the necessary modifications in the proofs being rather immediate. Accordingly, we shall consider *functions of squares*, and our references to Banach are to be interpreted in the sense that our statements are more or less obvious modifications of, but not necessarily equivalent to the corresponding statements of Banach.

2.2. We assume that the *function of squares* $F(s)$ is defined for all closed squares s in the fundamental square

$$S_0 : 0 \leqq u \leqq 1, \qquad 0 \leqq v \leqq 1.$$

$F(s)$ will be said to be of *bounded variation* if there exists a finite constant M such that

$$\sum_{j=1}^{m} | F(s_j) | \leqq M$$

for every system of closed squares without common interior points (Banach [1]).

2.3. $F(s)$ will be said to be *absolutely continuous* if there exists for every $\epsilon > 0$ an $\eta(\epsilon) > 0$ such that

$$\sum_{j=1}^{m} | F(s_j) | \leqq \epsilon$$

for every system of squares, without common interior points, which satisfy the condition[6]

$$\sum_{j=1}^{m} | s_j | \leqq \eta(\epsilon)$$

[5] The function of squares $\Gamma(s)$ of §8.9 comes under this description.

[6] Cf. footnote 1.

(Banach [1]). It follows immediately that the same inequality holds then also for *infinite* sequences of squares. This remark applies also to the inequality in §2.2.

2.4. It is easily seen that $F(s)$ may be absolutely continuous without being of bounded variation. If however $F(s)$ is *bounded*, that is, if $|F(s)|$ is less than a finite constant K independent of s, then absolute continuity implies that $F(s)$ is of bounded variation.[7] Indeed, consider the $\eta(\epsilon)$ appearing in the definition of absolute continuity, and put $2\delta = \eta(\epsilon)$. Take a system $s_1, \cdots, s_m$ of closed squares without common interior points, and assume first that $|s_j| < \delta$ $(j = 1, \cdots, m)$. Divide the system of squares into groups $\mathfrak{G}_1, \cdots, \mathfrak{G}_k$ in such a way that in each group the sum of the areas is less than 2δ. This can be done on account of our present assumption, and we have a grouping where the number k of groups is a *minimum*. In this extremal grouping there is at most *one* group whose squares have a total area less than δ, since otherwise two of the groups could be thrown together. Hence

$$(k-1)\delta \leqq \sum_{j=1}^{m} |s_j| \leqq |S_0| = 1,$$

and therefore

$$k \leqq 1 + \frac{1}{\delta}.$$

But the sum $\sum |F(s)|$, extended over the squares of any one group, is $\leqq 1$ by the definition of δ. Hence

$$\sum_{j=1}^{m} |F(s_j)| \leqq k \leqq 1 + \frac{1}{\delta}.$$

Assume, in the second place, that $|s_j| \geqq \delta$ $(j = 1, \cdots, m)$. Then

$$m\delta \leqq \sum_{j=1}^{m} |s_j| \leqq |S_0| \leqq 1,$$

and hence

$$\sum_{j=1}^{m} |F(s_j)| \leqq mK \leqq \frac{K}{\delta}.$$

Clearly, we have then for any system of closed squares $s_1, \cdots, s_m$ without common interior points, as a consequence of the preceding remarks, the inequality

$$\sum_{j=1}^{m} |F(s_j)| \leqq 1 + \frac{K+1}{\delta},$$

which shows that $F(s)$ is of bounded variation.

[7] Cf. Saks [1], p. 93 for the corresponding one-dimensional theorem.

2.5. Since the boundedness of $F(s)$ was used only to control *large* squares, it follows that if $F(s)$ is absolutely continuous, then there exist two positive constants M and δ, such that we have for every system of squares $s_1, \cdots, s_m$ without common interior points the inequality

$$\sum_{j=1}^{m} | F(s_j) | \leqq M,$$

provided that each square of the system has an area less than δ.

2.6. Take a point (u, v) in S_0, and consider all possible sequences of closed squares s_j such that (i) s_j contains (u, v), (ii) s_j is contained in S_0, and (iii) $| s_j | \to 0$. The least upper bound of lim sup $F(s_j)/| s_j |$, for all such sequences s_j, is the *upper derivative* $\bar{F}'(u, v)$ of $F(s)$ at (u, v). The *lower derivative* $\underline{F}'(u, v)$ is defined in a similar way. These derivatives may be equal to $\pm \infty$. They are always measurable (Banach, [1]). If $\bar{F}'(u, v)$, $\underline{F}'(u, v)$ are *finite* and equal at a point (u, v), then their common value is denoted by $F'(u, v)$ and is called the *derivative* of $F(s)$ at (u, v). If $F'(u, v)$ exists almost everywhere in S_0, then it is measurable there.

2.7. We shall use O to denote a relatively open subset of S_0, that is, a set which is the product of S_0 and of an open set. C will denote a closed subset of S_0. The subdivision of S_0 into j^2 congruent squares will be denoted by D_j, the class of those closed squares s of D_j which are contained in O by $K_j(O)$, and the class of those closed squares of D_j which have some point in common with C by $K_j(C)$.

2.8. With these notations we have the following theorem (Banach [1]). *If $F(s)$ is absolutely continuous and if $F'(u, v)$ exists almost everywhere in S_0, then $F'(u, v)$ is summable and*

$$\sum_{s \in K_j(O)} F(s) \xrightarrow[j\to\infty]{} \iint_O F'(u, v)\, dudv, \qquad \sum_{s \in K_j(C)} F(s) \xrightarrow[j\to\infty]{} \iint_C F'(u, v)\, dudv.$$

2.9. Suppose $F(s)$ satisfies the following conditions.

(i) $F(s) \geqq 0$.

(ii) If $s_1, \cdots, s_m$ are closed squares without common interior points, and if s is any closed square in S_0 which contains $s_1, \cdots, s_m$, then always

$$\sum_{j=1}^{m} F(s_j) \leqq F(s).$$

Then $F'(u, v)$ exists almost everywhere and is summable in S_0. While this theorem is not explicitly stated by Banach, it follows immediately by his reasoning in Banach [1].

2.10. Consider a closed square s in S_0. By a *decomposition* $D(s)$ of s we shall mean any finite or infinite sequence $s_1, \cdots, s_j, \cdots$ of closed squares with the following properties. (i) The squares of the sequence have no common interior points. (ii) $s = \sum s_j$ $(j = 1, 2, \cdots)$. If we use subdivisions of s into congruent squares, the following statement is easily proved. Given, in s, a finite system of closed squares $s_1, \cdots, s_m$ without common interior points, and given an $\epsilon > 0$, there exists a sequence $s_{m+1}, s_{m+2}, \cdots$ of closed squares with diameters less than ϵ such that $s_1, \cdots, s_m, s_{m+1}, \cdots$ is a decomposition of s. If the given squares $s_1, \cdots, s_m$ constitute by themselves a decomposition, then the assertion is vacuously true.

2.11. $F(s)$ will be called *normal* if (i) $F(s)$ is of bounded variation and (ii) we have $F(s) \leqq \sum F(s_j)$ for every decomposition $s_1, \cdots, s_j, \cdots$ of s (cf. Banach [1]).

2.12. *If $F(s)$ is normal, then $F'(u, v)$ exists almost everywhere and is summable in S_0* (Banach [1]).

3. On the topological index[8]

3.1. It will be convenient to use complex numbers $w = u + iv$, $z = x + iy$ in the sequel. Let there be given, on a Jordan curve C in the w-plane, a continuous function $z = f(w)$. If w describes C in the counter-clockwise sense, then z describes in the z-plane a *directed closed continuous curve* $\bar{C}$, and if z_0 is a point not on $\bar{C}$, then the change of the continuously varying argument of $z - z_0 = f(w) - z_0$ is of the form $2n_0\pi$, where n_0 is an integer (positive, negative or zero). This integer n_0 is called *the topological index* of the point z_0 with respect to $\bar{C}$.

3.2. Suppose $f(w)$ is continuous in the closed region bounded by a Jordan curve C. Let w_0 be a point interior to C, and put $z_0 = f(w_0)$. Assume that $f(w) \neq z_0$ in and on C, except at w_0. Let $\bar{C}$ and n_0 have the same meaning as in §3.1. *Then there exists a $\rho > 0$ such that for $0 < |z - z_0| < \rho$ the equation $f(w) = z$ has at least $|n_0|$ distinct roots in the interior of C.* By simple transformations, this statement can be reduced to the special case $w_0 = 0$, $z_0 = 0$ considered in Radó [2].

3.3. Consider again a function $f(w)$, continuous in the closed region bounded by a Jordan curve C. Let $\bar{C}$ have the same meaning as in §3.1. Take in the z-plane a point z_0 not on $\bar{C}$, and denote by $M_C(f(w) - z_0)$ the maximum of

[8] Cf. footnote 2. The notion of the topological index can be extended to spaces of higher dimension. With the exception of the lemma of 3.2, the facts stated in this chapter are special cases of well-known n-dimensional theorems. See also for further references Hopf [1].

$| f(w) - z_0 |$ on C. Assume that the topological index of z_0 with respect to $\bar{C}$ is equal to zero. *Then there exists a function $h(w)$ with the following properties.*

(i) *$h(w)$ is continuous in the closed region bounded by C.*

(ii) *$h(w) = f(w)$ on C.*

(iii) *$M(h(w) - z_0) = M_C(f(w) - z_0)$, where $M(h(w) - z_0)$ denotes the maximum of $| h(w) - z_0 |$ in the closed region bounded by C.*

(iv) *$h(w) \neq z_0$ in the closed region bounded by C.*

Proof. Since the topological index of z_0 with respect to $\bar{C}$ is equal to zero, we can write

$$f(w) - z_0 = | f(w) - z_0 | (\cos \varphi(w) + i \sin \varphi(w)), \qquad w \text{ on } C,$$

where $\varphi(w)$ is *single-valued*, real and continuous on C. Since $f(w) \neq z_0$ on C, the (real) logarithm of $| f(w) - z_0 |$ is continuous on C. Let $\xi(w)$, $\eta(w)$ denote the solutions of the Dirichlet problem, in the closed region bounded by C, for the boundary conditions

$$\left.\begin{aligned} \xi(w) &= \log | f(w) - z_0 | \\ \eta(w) &= \varphi(w) \end{aligned}\right\} \text{on } C.$$[9]

Define, in the region bounded by C,

$$h(w) = z_0 + \exp(\xi(w) + i\eta(w)).$$

Then $h(w)$ is single-valued and continuous in and on C, and it is clearly different from z_0 there. Obviously $h(w) = f(w)$ on C. Finally we have

$$| h(w) - z_0 | = \exp \xi(w).$$

As $\xi(w)$ is harmonic, it follows that $| h(w) - z_0 |$ takes on its maximum on the boundary, and as $h(w) = f(w)$ on C, it is proved that

$$M(h(w) - z_0) = M_C(f(w) - z_0).$$

3.4. For easier reference we state the following well-known theorem. Let $\mathfrak{R}$ be a closed bounded region in the w-plane, bounded by a finite number of Jordan curves $C_1, \cdots, C_m$. Consider a function $f(w)$ which is continuous and different from zero in $\mathfrak{R}$. Denote by V_j the change on C_j of the continuously varying argument of $f(w)$, the curve C_j being described in the positive direction relative to $\mathfrak{R}$ (that is, the exterior boundary curve is described in the counter-clockwise sense, while the interior boundary curves, if any, are described in the clockwise sense). *Then* $V_1 + \cdots + V_m = 0$.

[9] The use of harmonic functions could be avoided here, but the proof would be a little longer.

4. The function[10] $N_E(z)$

4.1. Throughout this paper T will denote a continuous transformation given by equations

$$T\colon x = x(u, v), \qquad y = y(u, v),$$

where $x(u, v)$, $y(u, v)$ are continuous in the closed square

$$S_0\colon 0 \leqq u \leqq 1, \qquad 0 \leqq v \leqq 1.$$

We shall use also the complex notation $w = u + iv$, $z = x + iy$. Then T is given by an equation

$$T\colon z = f(w),$$

where $f(w)$ is continuous in S_0. The perimeter of S_0 will be denoted by B_0. If w describes B_0 in the counter-clockwise sense, then its image z describes a directed continuous curve which we shall denote by $\bar{B}_0$. If E is a subset of S_0, and z a point in the z-plane, then $N_E(z)$ will denote *the number of distinct models* of z in E (that is, the number of distinct solutions of the equation $z = f(w)$ in the set E). If the number of models is infinite, then we put $N_E(z) = \infty$. To simplify the formulas, we shall write $N(z)$ for $N_{S_0}(z)$. The set of points z for which $N(z) = \infty$ will be denoted by $\mathfrak{S}_\infty$.

4.2. Take a set Q in S_0, and denote by $\bar{Q}$ its image in the z-plane. Take then a set $\bar{E}$ in the z-plane, and denote by E its *complete model* in S_0 (that is, E is the set of all points w in S_0 whose image is in $\bar{E}$). We have then the obvious relation

$$N_{EQ}(z) = \begin{cases} N_Q(z) \text{ for } z \text{ in } \bar{E}, \\ 0 \text{ for } z \text{ not in } \bar{E}. \end{cases}$$

4.3. We shall use the symbols O, C, D_j, $K_j(O)$, $K_j(C)$ in the same sense as in §2.7. It will be convenient to use subdivisions D_{p_j}, where p_j is the j-th positive prime. The following obvious remark will prove useful in the sequel. Given any point w_0 in the closed square S_0, there exists a $\delta = \delta(w_0) > 0$ such that for $j > \delta$ the point w_0 is comprised in *exactly one* closed square of the subdivision D_{p_j}. If w_0 is an interior point of S_0, then for $j > \delta$ the point w_0 will be *interior* to some square of D_{p_j}.

4.4. If s is a closed square in S_0, then $g_s(z)$ will denote the characteristic function[11] of its image in the z-plane. As the image of s is a closed set, $g_s(z)$ is measurable. The following four lemmas (§§4.5–4.8) are obvious consequences of the definitions and of the remarks in §4.3 concerning D_{p_j}.

[10] This function was introduced in Banach [2].

[11] The characteristic function of a point-set is equal to one at points of the set and equal to zero otherwise.

4.5. For every relatively open set O in S_0 we have

$$\sum_{s \in K_{p_j}(O)} g_s(z) \xrightarrow[j\to\infty]{} N_O(z). \tag{3}$$

4.6. For every closed set C in S_0 we have

$$\sum_{s \in K_{p_j}(C)} g_s(z) \xrightarrow[j\to\infty]{} N_C(z) \quad \text{for } z \text{ not in } \mathfrak{S}_\infty.$$

4.7. If $E, E_1, \cdots, E_j, \cdots$ are sets in S_0 such that

$$E = \prod_{j=1}^{\infty} E_j, \qquad E_1 \supset \cdots \supset E_j \supset \cdots,$$

then

$$N_{E_j}(z) \searrow N_E(z) \quad \text{for } z \text{ not in } \mathfrak{S}_\infty.$$

4.8. If $E, E_1, \cdots, E_j, \cdots$ are sets in S_0 such that

$$E = \sum_{j=1}^{\infty} E_j, \qquad E_1 \subset \cdots \subset E_j \subset \cdots,$$

then $N_{E_j}(z) \nearrow N_E(z)$.

4.9. We mention a few immediate consequences of the preceding remarks. For every j, the left side of (3) is a measurable function of z. Hence (3) shows that $N_O(z)$ is a measurable function of z. In particular $N(z)$ is measurable, and thus the set $\mathfrak{S}_\infty$ is also measurable. Similarly, it follows from §4.6 that $N_C(z)$ is measurable on the complement of $\mathfrak{S}_\infty$. If $\mathfrak{S}_\infty$ happens to be of measure zero, then it follows that $N_C(z)$ is also measurable for every closed set C in S_0.

4.10. Take a *measurable* set E in S_0. Then we can represent E in the form

$$E = e + \Gamma, \qquad e\Gamma = 0, \qquad |e| = 0, \qquad \Gamma = \sum_{j=1}^{\infty} C_j,$$

where $C_1, \cdots, C_j, \cdots$ are closed sets such that $C_1 \subset \cdots \subset C_j \subset \cdots$. By §4.8 we have then

$$N_E(z) = N_e(z) + \lim_{j\to\infty} N_{C_j}(z).$$

Let $\bar{e}$ be the image of e. Then $N_e(z) = 0$ for z not in $\bar{e}$. Hence by the final remark in §4.9 it follows that if T carries sets of measure zero into sets of measure zero, and if the set $\mathfrak{S}_\infty$ is of measure zero, then $N_E(z)$ is measurable for every measurable set E in S_0, and T carries measurable sets into measurable sets. The last assertion follows by the remark that the image of E is identical with the set where $N_E(z) \geqq 1$.

4.11. If s is a closed square in S_0, then $\mathfrak{E}_s$ will denote the set of those points z_0 for which there exists a $\delta = \delta(z_0) > 0$ such that $0 < |z - z_0| < \delta$ implies

$N_s(z) > N_s(z_0)$. Then it is easily seen that $\mathfrak{E}_s$ is a *denumerable* set (Radó [2]). Note that the definition implies that each point of $\mathfrak{E}_s$ has at most a finite number of models in s.

5. Auxiliary functions defined in terms of the topological index

5.1. The continuous transformation T being given as in §4.1, two points w_1, w_2 will be called *relatives* if they have the same image. The set of those *interior* points of S_0 which have a neighborhood clear of relatives will be denoted by I. Consider a point w_0 of I. Take a Jordan curve C_0 in S_0 which contains w_0 in its interior. If a point w describes C_0 in the counter-clockwise sense, then its image z describes a directed closed continuous curve $\bar{C}_0$. If the diameter of C_0 is small, then the image z_0 of w_0 will not be on $\bar{C}_0$, and the topological index of z_0 with respect to $\bar{C}_0$ will be independent of the particular choice of C_0 by §3.4. This index depends therefore only upon w_0. We shall denote it by $i(w_0)$. Dropping the subscript zero, we define in this manner a function $i(w)$ on I. We put $i(w) = 0$ for w in $S_0 - I$. Then $i(w)$ is defined for every point in S_0. This function $i(w)$ was introduced by Schauder [1]. For every integer k we define a set I_k as the subset of I where $i(w) = k$. Note that I_0 is the set such that (i) $i(w) = 0$, (ii) w is *interior* to S_0, and (iii) w has some neighborhood clear of relatives. Finally we denote by I^* the set where $|i(w)| > 1$ and by $\bar{I}$, $\bar{I}_k$, $\bar{I}^*$ the images of the sets I, I_k, I^*, respectively. The sets I, I_k, I^* are measurable for every continuous transformation T (Schauder [1]; for a somewhat different proof see Radó [2]).

Given a closed square s in S_0, we denote its perimeter by b. If a point describes b in the counter-clockwise sense, then its image in the z-plane describes a directed closed continuous curve $\bar{b}$. We define a function $n_s(z)$ as follows. For z not on $\bar{b}$, $n_s(z)$ is equal to the topological index of z with respect to $\bar{b}$. For z on $\bar{b}$, we put $n_s(z) = 0$. This function $n_s(z)$ is clearly measurable. To simplify the formulas, we shall write $n(z)$ for $n_{S_0}(z)$.

5.2. LEMMA. *For $|k| > 1$ the set $\bar{I}_k$ is denumerable.*

Proof. Consider the subdivisions D_{p_j} of S_0. For every closed square s of D_{p_j}, consider the set $\mathfrak{E}_s$ of §4.11, and put

$$\sum_{j=1}^{\infty}\Big(\sum_{s\,\epsilon\,D_{p_j}}\mathfrak{E}_s\Big) = \mathfrak{S}.$$

By §4.11 the set $\mathfrak{S}$ is denumerable. Hence it is sufficient to show that $\bar{I}_k \subset \mathfrak{S}$ for $|k| > 1$. Take any point z_0 in $\bar{I}_k$, $|k| > 1$. Then z_0 has some model w_0 in I_k. For j large enough, w_0 will be interior to a square s_0 of D_{p_j} (see §4.3) and s_0 will contain no relative of w_0. The topological index of z_0, with respect to the image of the perimeter of s_0, will then be equal to $i(w_0) = k$. By §3.2, we shall have then $N_{s_0}(z) \geqq |k| > 1$ for z close to z_0, while $N_{s_0}(z_0) = 1$. That is, z_0 is in the set $\mathfrak{E}_{s_0}$ and hence in $\mathfrak{S}$.

5.3. Since $\bar{I}^* = \sum \bar{I}_k$, $|k| > 1$, we have the corollary: *The image of the set where $i(w)$ is different from 0, +1, −1 is a denumerable set, for every continuous transformation T.*

As an immediate consequence, we have the

Lemma (added February 27, 1938). *The set I^* is also denumerable. In other words, we have $|i(w)| \leqq 1$, except possibly on a denumerable set.*

Proof. Suppose I^* is not empty. By the preceding corollary, $\bar{I}^*$ is a finite or infinite sequence of points $z_1, \cdots, z_j, \cdots$. Denote by E_j the set of the models of z_j in I^*. Then $I^* = \sum E_j$, and it is sufficient to show that E_j is denumerable. Let w_j be any point of E_j. As $E_j \subset I^* \subset I$, the point w_j has a neighborhood clear of relatives and hence clear of further points of E_j, since *all* the points of E_j are relatives of w_j. Thus E_j is an isolated and hence denumerable set.

We leave it to the reader to formulate the obvious implications of this lemma in connection with §§ 7.9, 7.10, 7.11, 8.6, 9.1.

5.4. Suppose the point z_0 is not in the set $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$. Then z_0 has in S_0 a finite number of models

$$w_1, \cdots, w_m, \qquad m = N(z_0),$$

all of which are in the set I, but none of which is in the set I^*. Hence all these models are in $I_0 + I_{+1} + I_{-1}$, and thus

$$i(w_1) + \cdots + i(w_m) = N_{I_{+1}}(z) - N_{I_{-1}}(z). \tag{4}$$

Take, for $h = 1, \cdots, m$, a Jordan curve C_h contained in S_0 and containing w_h in its interior. If the diameter of C_h is small, then clearly (i) $i(w_h)$ is equal to the topological index of z_0 with respect to the image $\bar{C}_h$ of C (cf. §5.1), (ii) the curves $C_1, \cdots, C_m$ are exterior to each other, and (iii) the closed region bounded by $B_0, C_1, \cdots, C_m$ contains no model of z_0. Hence, by §3.4,

$$i(w_1) + \cdots + i(w_m) = n(z_0). \tag{5}$$

From (4) and (5) we infer that

$$n(z) = N_{I_{+1}}(z) - N_{I_{-1}}(z) \text{ for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0. \tag{6}$$

5.5. It follows from (6) that

$$|n(z)| \leqq N_{I_{+1}}(z) + N_{I_{-1}}(z) \leqq N(z) \text{ for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0. \tag{7}$$

But $n(z) = 0$ on $\bar{B}_0$, and $N(z) = \infty$ on $\mathfrak{S}_\infty$. Hence (7) holds on $\mathfrak{S}_\infty + \bar{B}_0$. We have therefore the theorem: *We have $|n(z)| \leqq N(z)$, except possibly on a denumerable set.* In a somewhat different fashion this theorem was proved in Radó [2].

5.6. Given a closed square s in S_0, the set where $n_s(z) \neq 0$ is open. The characteristic function $\gamma_s(z)$ of this set is therefore measurable. Consider now the subdivision D_{p_j} of S_0, and define

$$\psi_j(z) = \sum_{s \in D_{p_j}} \gamma_s(z).$$

We assert that

$$(8) \qquad \left.\begin{aligned} &\psi_j(z) \leqq N_{I_{+1}}(z) + N_{I_{-1}}(z) \\ &\psi_j(z) \xrightarrow[j\to\infty]{} N_{I_{+1}}(z) + N_{I_{-1}}(z) \end{aligned}\right\} \text{ for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0.$$

Proof. The assumptions imply that z has a finite number of models $w_1, \cdots, w_m$, all of which are in the set $I_0 + I_{+1} + I_{-1}$. Let s be a closed square of D_{p_j}. If there is some model of z on the perimeter of s, then $n_s(z) = 0$ and hence $\gamma_s(z) = 0$. If there is no model of z on the perimeter of s, then by §3.4 and by the definition of $i(w)$ we have $n_s(z) = \sum i(w_h)$, the summation being extended over those models w_h of z which are interior to s. Hence, only those squares s of D_{p_j} can contribute to $\psi_j(z)$ which contain in their interior some model w_h such that $i(w_h) \neq 0$. The number of such models, under the assumptions made concerning z, is equal to

$$N_{I_{+1}}(z) + N_{I_{-1}}(z).$$

As each contributing square of D_{p_j} contains at least one of these models in its interior, the inequality in formula (8) follows.

Keeping z fixed, let us increase j. By §4.3, for j exceeding a certain j_0 no square of D_{p_j} will contain more than one of the models $w_1, \cdots, w_m$ of z and each of these models will be *interior* to a square of D_{p_j}. Take $j > j_0$ and denote by s_h that square of D_{p_j} which contains w_h (and no other model of z) in its interior ($h = 1, \cdots, m$). We have then, by the definition of $i(w)$,

$$n_{s_h}(z) = i(w_h).$$

As w_h is in the set $I_0 + I_{+1} + I_{-1}$, it follows that $n_{s_h}(z)$ and hence $\gamma_{s_h}(z)$ is different from zero if and only if w_h is in $I_{+1} + I_{-1}$. That is, we have

$$(9) \qquad \sum_{h=1}^{m} \gamma_{s_h}(z) = N_{I_{+1}}(z) + N_{I_{-1}}(z).$$

On the other hand, we have

$$(10) \qquad \sum_{s \in D_{p_j}} \gamma_s(z) = \sum_{h=1}^{m} \gamma_{s_h}(z).$$

Indeed, a square s of D_{p_j} which is different from $s_1, \cdots, s_m$ contains no model of z and therefore $n_s(z) = 0$ and hence $\gamma_s(z) = 0$ for such a square. (9) and (10) imply that

$$\psi_j(z) = N_{I_{+1}}(z) + N_{I_{-1}}(z)$$

for large j, and the second relation (8) is also proved.

5.7. Take a closed square s in S_0 and consider a decomposition $D(s) : s_1, \cdots, s_j, \cdots$ in the sense of §2.10. The perimeter b_j of s_j is carried by T into a directed closed continuous curve $\bar{b}_j$. *We assert that*

$$\gamma_s(z) \leqq \sum_j \gamma_{s_j}(z) \qquad \text{for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0 + \sum_j \bar{b}_j. \tag{11}$$

Proof. The assumptions imply that z has a finite number of models $w_1, \cdots, w_m$, each of which is interior to some square of the decomposition $D(s)$. For simplicity, denote by $\sigma_1, \cdots, \sigma_h$ those squares of $D(s)$ which contain at least one model of z and by $\sigma_1^*, \cdots, \sigma_h^*$ squares obtained by slightly contracting $\sigma_1, \cdots, \sigma_h$, respectively. Then, by §3.4,

$$n_{\sigma_1}(z) = n_{\sigma_1^*}(z), \cdots, n_{\sigma_h}(z) = n_{\sigma_h^*}(z), \tag{12}$$

and

$$n_s(z) = n_{\sigma_1^*}(z) + \cdots + n_{\sigma_h^*}(z). \tag{13}$$

Suppose now first that $\gamma_s(z) = 1$. Then $n_s(z) \neq 0$, and from (12) and (13) it follows that at least one of $n_{\sigma_1}(z), \cdots, n_{\sigma_h}(z)$ and hence at least one of $\gamma_{\sigma_1}(z), \cdots, \gamma_{\sigma_h}(z)$ is different from zero, and thus (11) is obvious. Suppose secondly that $\gamma_s(z) = 0$. Then (11) is again obvious, since $\gamma_{s_j}(z) \geqq 0$.

6. The kernel

6.1. Given in S_0 two continuous transformations

$$T_1: z = f_1(w), \qquad T_2: z = f_2(w),$$

we define their *distance* $\| T_1, T_2 \|$ as the maximum of $| f_1(w) - f_2(w) |$ in the closed square S_0. If T_j, T are continuous transformations given in S_0, such that $\| T, T_j \| \to 0$, then we shall say that T_j converges to T.

6.2. Given a continuous transformation T as in §4.1, we define, for every non-negative integer k, a set $\mathfrak{K}_k$ in the z-plane as follows. A point z_0 belongs to $\mathfrak{K}_k$ if and only if there exists an $\epsilon = \epsilon(k, z_0) > 0$, such that for every continuous transformation

$$T^*: z = f^*(w), \qquad w \text{ in } S_0,$$

which satisfies the inequality $\| T, T^* \| < \epsilon$, the point z_0 has at least k distinct models in S_0 with respect to T^*. The set $\mathfrak{K}_0$ is then identical with the whole z-plane. Obviously, $\mathfrak{K}_1 \supset \mathfrak{K}_2 \supset \cdots$. We put

$$\mathfrak{K}_\infty = \prod_{k=0}^{\infty} \mathfrak{K}_k.$$

The set $\mathfrak{K}_k$ may be called *the kernel* of order k of the image of S_0. For $k = 1$, this concept was introduced by the author in a study of the area of continuous surfaces (Radó [1]) as a possible substitute for certain very complicated geo-

metrical concepts used by Geöcze. Observe that some or all of the sets $\mathfrak{K}_1, \cdots, \mathfrak{K}_k, \cdots$ may be empty.

6.3. *The sets $\mathfrak{K}_k$ are measurable for every continuous transformation T.* For $k = 1$, this was proved by Saks (see Radó [1]). For a general k the following simple reasoning may be used. For every positive integer m, define $\mathfrak{K}_{k,m}$ as the set of all points z_0 for which the following assertion is true: The point z_0 has at least k models in S_0 with respect to every continuous transformation T^* which satisfies the inequality $\| T, T^* \| < m^{-1}$. Clearly

$$\mathfrak{K}_k = \sum_{m=1}^{\infty} \mathfrak{K}_{k,m},$$

and therefore it is sufficient to show that $\mathfrak{K}_{k,m}$ is measurable. Consider a point z_0 which is *not* in $\mathfrak{K}_{k,m}$. Then we have a continuous transformation T^* such that

$$\| T, T^* \| < m^{-1}, \qquad N^*(z_0) < k, \tag{14}$$

where $N^*(z)$ denotes the number of models of z with respect to T^*. Put

$$\eta = m^{-1} - \| T, T^* \|.$$

Then $\eta > 0$. Take any point z_1 such that $| z - z_1 | < \eta$. Let

$$T\colon z = f(w), \qquad T^*\colon z = f^*(w), \qquad w \text{ in } S_0,$$

be the equations of T, T^*, respectively, and define a continuous transformation $\tilde{T}$ by the formula

$$\tilde{T}\colon z = f^*(w) + z_1 - z_0 .$$

Denote by $\tilde{N}(z)$ the number of models of z with respect to $\tilde{T}$. Clearly

$$\tilde{N}(z_1) = N^*(z_0), \tag{15}$$

and

$$\| T, \tilde{T} \| \leqq \| T, T^* \| + \| T^*, \tilde{T} \| < \| T, T^* \| + \eta = m^{-1}. \tag{16}$$

By (14), (15), (16) we have

$$\| T, \tilde{T} \| < m^{-1}, \qquad \tilde{N}(z_1) < k.$$

Hence z_1 is *not* a point of $\mathfrak{K}_{k,m}$. That is, for every point z_0 of the complement of $\mathfrak{K}_{k,m}$ there exists an $\eta > 0$ such that every point z_1, for which $| z_1 - z_0 | < \eta$, is also a point of the complement of $\mathfrak{K}_{k,m}$. In other words, the complement of $\mathfrak{K}_{k,m}$ is open, and hence $\mathfrak{K}_{k,m}$ is measurable.

6.4. Take a point z_0 such that z_0 is not in $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$. Then the location of z_0 with respect to the sets $\mathfrak{K}_k$ can be discussed as follows. The assumptions imply that z_0 has a finite number of models, all of which are in $I_0 + I_{+1}$

$+I_{-1}$. Denote by $w_1, \cdots, w_k$ those of the models which are in $I_{+1} + I_{-1}$, and by $w_{k+1}, \cdots, w_m$ those which are in I_0. For $j = 1, \cdots, k$, take a small circular disc d_j with center at w_j. The perimeter c_j of d_j is carried by T into a directed closed continuous curve $\bar{c}_j$. The topological index of z_0 with respect to $\bar{c}_j$ is then equal to $i(w_j) = \pm 1$. If we take a second continuous transformation

$$T^*: z = f^*(w), \qquad w \text{ in } S_0,$$

and if $\| T, T^* \|$ is sufficiently small, then the topological index of z_0 with respect to the image $\bar{c}_j^*$ of c_j under T^* will be equal to the topological index of z_0 with respect to $\bar{c}_j$, and hence it will be also equal to ± 1. If we apply the theorem of §3.4 to the function $f^*(w) - z_0$ in the disc d_j, then it follows that the point z_0 has at least one model in d_j with respect to T^*. That is, if $\| T, T^* \|$ is sufficiently small, then z_0 has at least k models with respect to T^*, namely, at least one in each of the discs $d_1, \cdots, d_k$. Hence z_0 is in $\mathfrak{K}_k$.

On the other hand, we shall see presently that z_0 is *not* in $\mathfrak{K}_{k+1}$. To prove this, we shall exhibit for any given $\epsilon > 0$ a continuous transformation T^* such that

$$\| T, T^* \| < \epsilon, \qquad N^*(z_0) < k + 1,$$

where $N^*(z)$ is the number of models of z with respect to T^*. Let $d_j, c_j, \bar{c}_j$ have the same meaning as before, except that this time we consider $j = k + 1, \cdots, m$. Again, for d_j small, the topological index of z_0 with respect to $\bar{c}_j$ is equal to $i(w_j)$ and hence equal to zero ($j = k + 1, \cdots, m$). Furthermore if d_j is small, then the maximum of $| f(w) - z_0 |$ on the disc d_j will be less than $\epsilon/2$. Hence, by §3.3, we have a function $h_j(w)$ with the following properties.

(i) $h_j(w)$ is continuous on d_j.

(ii) $h_j(w) = f(w)$ on c_j.

(iii) The maximum of $| h_j(w) - z_0 |$ on d_j is less than $\epsilon/2$.

(iv) $h_j(w) \neq z_0$ on d_j.

We define now a continuous transformation

$$T^*: z = f^*(w), \qquad w \text{ in } S_0,$$

as follows: $f^*(w) = h_j(w)$ on d_j ($j = k + 1, \cdots, m$) and $f^*(w) = f(w)$ otherwise in S_0. Then clearly $\| T, T^* \| < \epsilon$, while the models of z with respect to T^* are exactly the points $w_1, \cdots, w_k$. Hence $N^*(z_0) < k + 1$, and as ϵ was arbitrary, this proves that z_0 is not in $\mathfrak{K}_{k+1}$.

6.5. The result of §6.4 may be stated as follows. Take a point z not in $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$. Put

$$k = N_{I_{+1}}(z) + N_{I_{-1}}(z).$$

Then z is in $\mathfrak{K}_k$, but not in $\mathfrak{K}_{k+1}$.

6.6. Obviously, $N(z) \geqq k$ on $\mathfrak{K}_k$.

6.7. In S_0, take a sequence of continuous transformations T_j such that $T_j \to T$. Denote by $\mathfrak{R}_k^{(j)}$ the set defined in the same way in terms of T_j as $\mathfrak{R}_k$ was defined in terms of T. Suppose k is finite and take a point z in $\mathfrak{R}_k$. Then z is also in $\mathfrak{R}_k^{(j)}$ for sufficiently large j. The proof is obvious (cf. Radó [1] for the case $k = 1$).

6.8. We define now a function $\kappa(z)$ as follows. $\kappa(z) = k$ in $\mathfrak{R}_k - \mathfrak{R}_{k+1}$ if k is finite, and $\kappa(z) = +\infty$ in $\mathfrak{R}_\infty$. The preceding sections imply then the following facts concerning $\kappa(z)$.

6.9. $\kappa(z)$ is measurable, by §6.3.

6.10. $\kappa(z) \leqq N(z)$ by §6.6.

6.11. By §6.5 we have

$$\kappa(z) = N_{I_{+1}}(z) + N_{I_{-1}}(z) \text{ for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0. \tag{17}$$

6.12. Given in S_0 a sequence of continuous transformations T_j such that $T_j \to T$, let $\kappa_j(z)$ be defined in the same way in terms of T_j as $\kappa(z)$ was defined in terms of T. Then we have the relation

$$\kappa(z) \leqq \liminf_{j \to \infty} \kappa_j(z).$$

This follows immediately from §6.7.

6.13. Combining §6.11 and §5.6 we obtain the relations

$$\left.\begin{array}{l} \displaystyle\sum_{s \in D_{p_j}} \gamma_s(z) \leqq \kappa(z) \\ \displaystyle\sum_{s \in D_{p_j}} \gamma_s(z) \xrightarrow[j \to \infty]{} \kappa(z) \end{array}\right\} \text{ for } z \text{ not in } \mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0.$$

6.14. From §5.4 and §6.11 we infer that

$$| n(z) | = | N_{I_{+1}}(z) - N_{I_{-1}}(z) | \leqq N_{I_{+1}}(z) + N_{I_{-1}}(z) = \kappa(z),$$

for z not in $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$.

7. Questions of summability

7.1. The continuous transformation T being given as in §4.1, take a closed square s in S_0, denote by $g_s(z)$ the characteristic function of the image of s, and put

$$G(s) = \int\int g_s(z).$$

If the function of squares $G(s)$ is of bounded variation, in the sense of §2.2, then we shall say that T is BV (of bounded variation in the sense of Banach). If $G(s)$ is absolutely continuous, in the sense of §2.3, then we shall say that T is AC (absolutely continuous in the sense of Banach). Clearly, $G(s) \leqq G(S_0)$; that is, $G(s)$ is bounded. Hence, by §2.4, absolute continuity implies bounded variation in the present case. Obviously, $G(s)$ possesses the property (ii) of §2.11. Hence, if T is BV, then $G(s)$ is normal in the sense of §2.11. By §2.12, the derivative of $G(s)$ exists then almost everywhere in S_0, and it is summable there. We shall denote the derivative of $G(s)$ by $D(w)$.

7.2. If T is BV, then $N(z)$ is summable (Banach [2]). Indeed, by §4.5 we have (special case $O = S_0$)

$$\sum_{s \in D_{p_j}} g_s(z) \xrightarrow[j \to \infty]{} N(z). \tag{18}$$

Also, since T is BV,

$$\int\int_{S_0} \Big(\sum_{s \in D_{p_j}} g_s(z) \Big) = \sum_{s \in D_{p_j}} G(s) < M, \tag{19}$$

where M is a finite constant independent of the subdivision used. By a well-known theorem of Fatou,[12] (18) and (19) imply the summability of $N(z)$. As a corollary it follows that the set $\mathfrak{S}_\infty$ of §4.1 is of measure zero if T is BV (Banach [2]).

7.3. The following remark will prove useful. Take, in S_0, any system $s_1, \cdots, s_m$ of closed squares without common interior points. Then

$$\sum_{j=1}^{m} g_{s_j}(z) \leqq 4N(z). \tag{20}$$

Indeed, no model of z is contained in more than four of the closed squares $s_1, \cdots, s_m$.

7.4. Suppose T is such that the corresponding function $N(z)$ is summable. Integration of (20) yields then

$$\sum_{j=1}^{m} G(s_j) \leqq 4 \int\int N(z),$$

for every system of closed squares without common interior points. Hence T is BV. By §7.2 we have therefore the theorem: *T is BV if and only if $N(z)$ is summable* (Banach [2]).

[12] See Saks [1], p. 29.

7.5. Assume T is BV. For every relatively open set O (see §2.7) we have by §4.5 and §7.3:

$$\sum_{s \in K_{p_j}(O)} g_s(z) \xrightarrow[j\to\infty]{} N_O(z), \tag{21}$$

and

$$\sum_{s \in K_{p_j}(O)} g_s(z) \leqq 4N(z). \tag{22}$$

As $N(z)$ is summable, (22) implies by a well-known theorem of Lebesgue that term-wise integration of (21) is permissible. Hence

$$\sum_{s \in K_{p_j}(O)} G(s) \xrightarrow[j\to\infty]{} \int\int N_O(z). \tag{23}$$

7.6. If C is a closed set in S_0, then it follows in a similar fashion, by §§4.6, 7.2, 7.3, that

$$\sum_{s \in K_{p_j}(C)} G(s) \xrightarrow[j\to\infty]{} \int\int N_C(z),$$

if T is BV.

7.7. Assume T is BV. Take a relatively open subset O of S_0. Define a function $\varphi_j(w)$ as follows. If w is interior to a square s of $K_{p_j}(O)$, then $\varphi_j(w) = G(s)/|s|$. Otherwise $\varphi_j(w) = 0$. Then $\varphi_j(w) \to D(w)$ almost everywhere on O and $\varphi_j(w) \to 0$ almost everywhere on $S_0 - O$. Also

$$\int\int_O \varphi_j(w) = \sum_{s \in K_{p_j}(O)} G(s). \tag{24}$$

Hence, by the lemma of Fatou,

$$\int\int_O D(w) \leqq \liminf_{j\to\infty} \int\int_O \varphi_j(w),$$

and thus by (23) and (24)

$$\int\int_O D(w) \leqq \int\int N_O(z).$$

Similarly it follows that we have for every closed set C in S_0:

$$\int\int_C D(w) \leqq \int\int N_C(z),$$

if T is BV. In particular, for $O = S_0$, we have

$$\int\int_{S_0} D(w) \leqq \int\int N(z).$$

7.8. THEOREM. *Suppose T is BV. Let E be a subset of S_0 such that the image $\bar{E}$ of E is of measure zero. Then $D(w) = 0$ almost everywhere on E.*[13]

Proof. Since $N(z)$ is summable, we have for every positive integer j a $\delta_j > 0$ such that, if a measurable set $\mathfrak{S}$ in the z-plane satisfies the inequality $|\mathfrak{S}| < \delta_j$, then we have

$$\int\int_{\mathfrak{S}} N(z) < j^{-1}.$$

As $\bar{E}$ is of measure zero, we have an open set $\bar{O}_j$ such that

$$\bar{E} \subset \bar{O}_j, \qquad |\bar{O}_j| < \delta_j.$$

Denote by O_j the *complete model* of $\bar{O}_j$ (that is, the set of all points w whose image is in $\bar{O}_j$). Then O_j is clearly a relatively open subset of S_0, and $E \subset O_j$. Hence, by §7.7,

$$\int\int_{O_j} D(w) \leqq \int\int N_{O_j}(z) = \int\int_{\bar{O}_j} N(z) < j^{-1}.$$

Put

$$\Omega = \prod_{j=1}^{\infty} O_j.$$

Then

$$\int\int_{\Omega} D(w) \leqq \int\int_{O_j} D(w) < j^{-1}$$

for every positive integer j. Hence

$$\int\int_{\Omega} D(w) = 0,$$

and thus $D(w) = 0$ almost everywhere on Ω and *a fortiori* on the subset E of Ω.

7.9. By §5.3 it follows from §7.8 that if T is BV, then $D(w) = 0$ almost everywhere on the set I^* where $|i(w)| > 1$. Consider now the set $S_0 - I$. The perimeter B_0 of S_0 is of measure zero, and the image of $(S_0 - I) - B_0$ is comprised in the set $\mathfrak{S}_\infty$ which is of measure zero by §7.2. Hence, by §7.8, $D(w) = 0$ almost everywhere on $S_0 - I$. Summing up, we have the

THEOREM. *If T is BV, then $D(w) = 0$ almost everywhere on the set $S_0 - (I_0 + I_{+1} + I_{-1})$. Hence*

$$\int\int_{S_0} D(w) = \int\int_{I_0} D(w) + \int\int_{I_{+1}} D(w) + \int\int_{I_{-1}} D(w).$$

[13] Schauder proved this theorem under the assumption that T is AC (Schauder [1]).

7.10. Assuming that T is BV, we put $J(w) = i(w)D(w)$. This *generalized Jacobian* $J(w)$ was introduced by Schauder [1]. As $i(w)$ may be equal to any integer, it is interesting to note the

THEOREM. *If T is BV, then $|J(w)| \leqq D(w)$ almost everywhere in S_0.*

Indeed, on the set where $|J(w)| > D(w)$ we must have $|i(w)| > 1$ and $D(w) \neq 0$, and this set is of measure zero by §7.9.

7.11. In order to secure the summability of $J(w)$, Schauder [1] assumed that T is BV *and that $i(w)$ is bounded.* The author [2] proved then that $J(w)$ is summable if T is AC. As a consequence of §§7.10 and 7.1 we can state at present the

THEOREM. *If T is BV, then $J(w)$ is summable.*

7.12. If T is BV, then $J(w) = 0$ almost everywhere on the following sets: (i) on $S_0 - (I_0 + I_{+1} + I_{-1})$, since there $D(w) = 0$ almost everywhere by §7.9; (ii) on I_0, since there $i(w) = 0$. Thus the set where $J(w) \neq 0$ is a subset of $I_{+1} + I_{-1}$, if we disregard sets of measure zero. Hence we have the formulas

$$\int\int_{S_0} J(w) = \int\int_{I_{+1}} D(w) - \int\int_{I_{-1}} D(w),$$

$$\int\int_{S_0} |J(w)| = \int\int_{I_{+1}} D(w) + \int\int_{I_{-1}} D(w).$$

7.13. Let E be a measurable subset of S_0. A reasoning similar to that in §7.12 shows that

$$\int\int_{E} J(w) = \int\int_{E_+} D(w) - \int\int_{E_-} D(w),$$

$$\int\int_{E} |J(w)| = \int\int_{E_+} D(w) + \int\int_{E_-} D(w),$$

where $E_+ = I_{+1}E$, $E_- = I_{-1}E$.

7.14. Comparison of §§7.2 and 5.5 yields the theorem that *if T is BV, then $n(z)$ is summable.* This fact was already proved in an earlier paper of the author (Radó [2]).

7.15. Comparison of §§7.2 and 6.10 yields the theorem that *if T is BV, then $\kappa(z)$ is summable.*

8. Integral identities

8.1. Suppose the continuous transformation T, given as in §4.1, is AC (see §7.1). Take in S_0 a set e of measure zero. Denote by $\bar{e}$ the image of e. Given then any $\epsilon > 0$, take in S_0 a relatively open set O such that $e \subset O$ and

$\| O \| \leqq \eta(\epsilon)$, where $\eta(\epsilon)$ is the quantity appearing in the definition of absolute continuity (see §2.3). We can write

$$O = \sum_{j=1}^{\infty} s_j ,$$

where $s_1 , \cdots , s_j , \cdots$ are closed squares without common interior points. If $\bar{O}$, $\bar{s}_j$ are the images of O, s_j, respectively, then it follows that

$$| \bar{O} | \leqq \sum_{j=1}^{\infty} | \bar{s}_j | = \sum_{j=1}^{\infty} G(s_j) \leqq \epsilon,$$

since

$$\sum_{j=1}^{\infty} | s_j | = | O | \leqq \eta(\epsilon).$$

As $\bar{e}$ is a subset of $\bar{O}$ and ϵ is arbitrary, it follows that $\bar{e}$ is of measure zero. As T is also BV by §7.1, the set $\mathfrak{S}_\infty$ of §4.1 is of measure zero by §7.2. Hence, by §4.10, we have the theorem: *If T is AC, then sets of measure zero are carried into sets of measure zero, measurable sets are carried into measurable sets, and for every measurable set E in S_0 the function $N_E(z)$ is measurable* (Banach [2]).

8.2. Assuming that T is AC, we obtain from §§2.8, 7.5, 7.6 the formulas

$$\int\int_O D(w) = \int\int N_O(z), \qquad \int\int_C D(w) = \int\int N_C(z),$$

where O is any relatively open subset and C any closed subset of S_0. The special case $O = S_0$ is due to Banach [2] and both formulas are special cases of a theorem of Schauder which we shall derive presently from the preceding formulas.

8.3. Suppose T is AC, and take any measurable set E in S_0. We can write

$$E = e + \Gamma = e + \sum_{j=1}^{\infty} C_j ,$$

where $C_1 , C_2 , \cdots$ are closed sets and

$$| e | = 0, \qquad e\Gamma = 0, \qquad C_1 \subset C_2 \subset \cdots .$$

We have then

$$\int\int N_E(z) = \int\int N_e(z) + \int\int N_\Gamma(z) = \int\int N_\Gamma(z),$$

since $N_e(z) = 0$ for z not in the image $\bar{e}$ of e, and $\bar{e}$ is of measure zero by §8.1. Also, by §4.8,

$$N_{C_j}(z) \nearrow N_\Gamma(z),$$

and hence

$$\iint N_{C_j}(z) \to \iint N_\Gamma(z).$$

Obviously

$$\iint_{C_j} D(w) \to \iint_\Gamma D(w) = \iint_E D(w),$$

and by §8.2 we have

$$\iint_{C_j} D(w) = \iint N_{C_j}(z).$$

Combining these relations we obtain the

THEOREM. *If T is AC, then we have*

$$\iint_E D(w) = \iint N_E(z), \tag{25}$$

for every measurable set E in S_0 (Schauder [1]).

8.4. We have seen that if T is AC, then T is also BV and sets of measure zero are carried into sets of measure zero. Conversely, if T is BV and if sets of measure zero are carried into sets of measure zero, then T is AC (Banach [2], Schauder [1]). This statement is a special case of certain general results of Saks [2].

8.5. Since $N_E(z) = 0$ for z not in the image $\bar{E}$ of E, the theorems of §§8.3 and 7.8 imply that if T is AC, then the image $\bar{E}$ of a measurable set E is of measure zero if and only if $D(w) = 0$ almost everywhere on E. Schauder [1] contains a number of interesting applications of this fact.

8.6. If T is AC, then the set $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$ of §5.4 is of measure zero by §§7.2, 8.1, 5.3. Integration of formula (6) in §5.4 yields therefore

$$\iint n(z) = \iint N_{I_{+1}}(z) - \iint N_{I_{-1}}(z).$$

By §§8.3 and 7.12 this gives the

THEOREM. *If T is AC, then*

$$\iint_{S_0} J(w) = \iint n(z). \tag{26}$$

Schauder [1] proved this formula under the assumption that $i(w)$ is bounded. For the general case the formula was proved, in a less direct way, by the author (Radó [2]).

8.7. If T is AC, then integration of (17) in §6.11 yields

$$\int\int \kappa(z) = \int\int N_{I+1}(z) + \int\int N_{I-1}(z),$$

and by §§8.3 and 7.12 we obtain the

THEOREM. *If T is AC, then*

$$\int\int_{S_0} |J(w)| = \int\int \kappa(z).\text{[14]} \tag{27}$$

8.8. Consider now, in S_0, a sequence T_j of continuous transformations, such that $T_j \to T$ in the sense of §6.1. By §6.12 we have then

$$\kappa(z) \leqq \liminf_{j\to\infty} \kappa_j(z),$$

and therefore, by the lemma of Fatou,

$$\int\int \kappa(z) \leqq \liminf_{j\to\infty} \int\int \kappa_j(z),$$

provided only that the integrals involved exist. By §§6.10, 7.1, 7.2 this is certainly the case if the transformations involved are BV. If they are also AC, then §8.7 yields the

THEOREM. *If $T_j \to T$ in S_0, and if all these transformations are AC, then*

$$\int\int_{S_0} |J(w)| \leqq \liminf_{j\to\infty} \int\int_{S_0} |J_j(w)|. \tag{28}$$

In this statement, $J_j(w)$ has the same meaning with respect to T_j as that of $J(w)$ with respect to T. Formula (28) expresses the fact that the integral of $|J(w)|$ over S_0 is, in its dependence upon T, *a lower semi-continuous functional.*

8.9. We shall derive presently a new geometrical interpretation for $J(w)$ and $|J(w)|$. Assuming that T is AC, let us consider the function $\gamma_s(z)$ of §5.6. We put

$$\Gamma(s) = \int\int \gamma_s(z).\text{[15]}$$

That is, $\Gamma(s)$ is the measure of the set where $n_s(z) \neq 0$. From §3.4 it follows readily that this set is comprised in the image of s. Hence $\Gamma(s) \leqq G(s)$. By

[14] While formula (25), for $E = S_0$, and formula (26) give geometrical interpretations for the integrals of $D(w)$ and $J(w)$, respectively, formula (27) gives such an interpretation for the integral of the absolute value of the generalized Jacobian. The geometrical meaning of this last integral was not considered by Banach and Schauder. Formula (27) may be considered as an expression for the area of the *stable* portion of the image of the fundamental square S_0.

[15] The function of squares $\Gamma(s)$ has already been used in Radó [1].

assumption $G(s)$ is absolutely continuous, and therefore $\Gamma(s)$ is a fortiori absolutely continuous. As the set $\mathfrak{S}_\infty + \bar{I}^* + \bar{B}_0$ is at present of measure zero, we obtain by integrating formula (11) of §5.7 the inequality

$$\Gamma(s) \leqq \sum_j \Gamma(s_j)$$

for every decomposition $D(s): s_1, s_2, \cdots$ of every square s in S_0. Thus the function of squares $\Gamma(s)$ is *normal* in the sense of §2.11. By §2.12, its derivative $\Gamma'(w)$ exists almost everywhere and is summable in S_0. By §2.8 (special case $O = S_0$) we have therefore

$$\sum_{s \in D_{p_j}} \Gamma(s) \xrightarrow[j\to\infty]{} \int\int_{S_0} \Gamma'(w).$$

On the other hand, formulas (8) of §5.6 yield by integration the formula

$$\sum_{s \in D_{p_j}} \Gamma(s) \xrightarrow[j\to\infty]{} \int\int N_{I_{+1}}(z) + \int\int N_{I_{-1}}(z).$$

By §§8.3 and 7.12 we have

$$\int\int N_{I_{+1}}(z) + \int\int N_{I_{-1}}(z) = \int\int_{I_{+1}} D(w) + \int\int_{I_{-1}} D(w) = \int\int_{S_0} | J(w) |.$$

Hence

$$\int\int_{S_0} \Gamma'(w) = \int\int_{S_0} | J(w) |.$$

Observe now that T is AC in every square s in S_0, and that $\Gamma'(w)$, $J(w)$ depend only upon the local behavior of T. The preceding formula remains therefore valid if we replace S_0 by any square s in S_0. That is,

$$\int\int_s \Gamma'(w) = \int\int_s | J(w) |$$

for every square s in S_0. *Hence*

$$| J(w) | = \Gamma'(w) \tag{29}$$

almost everywhere in S_0. *We assert that we have also*

$$J(w) = i(w)\Gamma'(w) \tag{30}$$

almost everywhere in S_0. If $J(w) = 0$, then $\Gamma'(w) = 0$ by (29), and (30) is true. If $J(w) \neq 0$, then (cf. §7.12) w is in the set $I_{+1} + I_{-1}$ (if we disregard sets of measure zero), and therefore $J(w)$ has the same sign as $i(w) = \pm 1$. Thus (30) is obvious in this case also.

9. Transformation of double integrals

9.1. Assuming that the continuous transformation T, given as in §4.1, is AC, we shall discuss the range of validity of the formula of transformation

$$\int\int_{S_0} H(f(w))J(w) = \int\int H(z)n(z), \tag{31}$$

where $H(z)$ is a measurable function. Schauder discussed (31) under the additional assumption that $i(w)$ is bounded. Inspection of his work shows that he used this assumption only to secure the summability of $J(w)$ and of $n(z)$. Since we know that $J(w)$ and $n(z)$ are automatically summable as soon as T is BV (see §§7.11, 7.14), we are in a position to generalize the results of Schauder. Following his line of reasoning[16] we would obtain immediately the theorem that (31) holds whenever *both* integrals involved exist in the Lebesgue sense. We shall indicate briefly a different proof which leads to a more precise result. This proof is based upon the analogous formula

$$\int\int_{S_0} H(f(w)) \mid J(w) \mid = \int\int H(z)\kappa(z), \tag{32}$$

interesting in itself, which we shall establish simultaneously.

9.2. We shall prove the following

THEOREM. *If T is AC, and if $H(z)$ is measurable and finite,*[17] *then the four functions*

$$H(f(w)) \mid J(w) \mid, \qquad H(f(w))J(w), \qquad H(z)\kappa(z), \qquad H(z)n(z) \tag{33}$$

are measurable. If any one of the first three functions is summable, then all four functions are summable, and the formulas (31) *and* (32) *are both valid.*

Remark. The example

$$T: \begin{cases} x = u,\ y = v & \text{for } 0 \leqq u \leqq \frac{1}{2},\ 0 \leqq v \leqq 1, \\ x = 1 - u,\ y = v & \text{for } \frac{1}{2} \leqq u \leqq 1,\ 0 \leqq v \leqq 1 \end{cases}$$

shows that the existence, in the Lebesgue sense, of the right side of (31) does not imply generally the existence of the left side.

9.3. We proceed to prove the various assertions of the preceding theorem. In the z-plane, consider a set of finite functions $H_j(z)$, $H(z)$ such that $H_j(z) \to H(z)$ almost everywhere. We assert that

$$H_j(f(w)) \mid J(w) \mid \to H(f(w)) \mid J(w) \mid,\ H_j(f(w))J(w) \to H(f(w))J(w) \tag{34}$$

[16] In Schauder [1].

[17] That is, $H(z) \neq \pm\infty$. The assumption of finiteness does not imply boundedness and is made only to avoid trivial discussions.

almost everywhere in S_0. Indeed, denote by $\bar{E}$ the set on which the relation $H_j(z) \to H(z)$ does not hold. By assumption $|\bar{E}| = 0$. Let E be the *complete model* of $\bar{E}$.[18] On $S_0 - E$ we have $H_j(f(w)) \to H(f(w))$, and (34) is obviously true on this set. By §7.8, we have $D(w) = 0$ and hence $J(w) = 0$ almost everywhere on E, and thus (34) holds almost everywhere in S_0.

The same kind of reasoning shows that if we are given two finite functions $H_1(z)$, $H_2(z)$ such that $H_1(z) = H_2(z)$ almost everywhere in the z-plane, then we have also

$$H_1(f(w)) \mid J(w) \mid = H_2(f(w)) \mid J(w) \mid, \qquad H_1(f(w))J(w) = H_2(f(w))J(w)$$

almost everywhere in S_0.[19]

9.4. Suppose now that $H(z)$ is finite and measurable in the z-plane. We have then a sequence $H_j(z)$ of continuous functions such that $H_j(z) \to H(z)$ almost everywhere. The functions $H_j(f(w))$ being continuous, it follows by §§2.6, 5.1, 6.8, 9.3 that the functions (33) are measurable.

9.5. Denote by C^* the class of all those finite measurable functions $H(z)$ for which the formula (32) holds. We shall show that C^* contains certain simple functions, and we shall discuss then (32) by means of the following remarks.

9.6. If $H_1(z), \cdots, H_m(z)$ are in C^*, and if $a_1, \cdots, a_m$ are constants, then $a_1H_1(z) + \cdots + a_mH_m(z)$ is also in C^*. The proof is obvious.

9.7. Suppose $H_j(z)$, $H(z)$ are finite measurable functions, such that $H_j(z) \to H(z)$ almost everywhere. Suppose also that $H_j(z)$ is in C^* $(j = 1, 2, \cdots)$. By §9.3 we have then

$$H_j(f(w)) \mid J(w) \mid \to H(f(w)) \mid J(w) \mid \tag{35}$$

almost everywhere in S_0, while by assumption

$$H_j(z)\kappa(z) \to H(z)\kappa(z) \tag{36}$$

almost everywhere in the z-plane, and

$$\int\int_{S_0} H_j(f(w)) \mid J(w) \mid = \int\int H_j(z)\kappa(z)$$

for $j = 1, 2, \cdots$. Hence $H(z)$ will belong to C^* whenever it is permissible to integrate term by term the relations (35) and (36). By well-known theorems on term-wise integration,[20] this will be clearly the case under the following circumstances (note that $\mid J(w) \mid$ and $\kappa(z)$ are summable and non-negative).

[18] That is, E is the set of all those points w in S_0 whose image is in $\bar{E}$.

[19] It follows that if we change the values of a function $H(z)$ on a set of measure zero, then the values of the integrands on the left sides of (31) and (32) are changed at most on a set of measure zero.

[20] Cf. Saks [1], Chapter 1.

Case (a). $| H_j(z) | < M$, where M is a finite constant independent of j.

Case (b). $H(z)\kappa(z)$ and $H(f(w)) \mid J(w) \mid$ are summable, and $| H_j(z) | \leqq | H(z) |$ almost everywhere.

Case (c). $H_j(z) \geqq 0$, $H_j(z) \nearrow H(z)$ almost everywhere, and one of the functions $H(f(w)) \mid J(w) \mid$, $H(z)\kappa(z)$ is summable.

9.8. Suppose now that $H(z)$ is the characteristic function of a bounded open set $\bar{O}$ in the z-plane. Denote by O the *complete model* of $\bar{O}$. Then O is clearly open relative to S_0, and $H(f(w))$ is the characteristic function of O. By §6.11 we have

$$\kappa(z) = N_{I_{+1}}(z) + N_{I_{-1}}(z)$$

almost everywhere. By integration we obtain, with regard to §§4.2, 8.3, 7.13,

$$\begin{aligned}\int\int H(z)\kappa(z) &= \int\int_{\bar{O}} \kappa(z) = \int\int_{\bar{O}} N_{I_{+1}}(z) + \int\int_{\bar{O}} N_{I_{-1}}(z) \\ &= \int\int N_{OI_{+1}}(z) + \int\int N_{OI_{-1}}(z) = \int\int_{OI_{+1}} D(w) + \int\int_{OI_{-1}} D(w) \\ &= \int\int_{O} \mid J(w) \mid = \int\int_{S_0} H(f(w)) \mid J(w) \mid .\end{aligned}$$

That is, the characteristic functions of bounded open sets in the z-plane belong to the class C^*.

9.9. By §§9.6, 9.8, 9.7(a) it follows that every bounded measurable function $H(z)$ belongs to C^*.

9.10. By §§9.9 and 9.7(b) it follows further that if $H(z)$ is measurable and finite, and if $H(f(w)) \mid J(w) \mid$ and $H(z)\kappa(z)$ are both summable, then $H(z)$ belongs to C^*.

9.11. From §9.6(c) and §9.9 we infer that if $H(z)$ is measurable, finite, and non-negative, and if one of the functions $H(f(w)) \mid J(w) \mid$, $H(z)\kappa(z)$ is summable, then $H(z)$ is in C^*.

9.12. To discuss the general case, let us use the notations

$$r^+ = \max(r, 0), \qquad r^- = \min(r, 0),$$

where r is a real number. If r is a function, then r^+, r^- are called the positive part and the negative part of r, respectively. Consider then a finite measurable function $H(z)$. As $| J(w) | \geqq 0$, $\kappa(z) \geqq 0$, we have

$$[H(f(w)) \mid J(w) \mid]^{\pm} = [H(f(w))]^{\pm} \mid J(w) \mid, \qquad [H(z)\kappa(z)]^{\pm} = [H(z)]^{\pm}\kappa(z).$$

If we observe that a function is summable if and only if its positive part and its negative part are both summable, it follows immediately from §9.11 that

if $H(z)$ is finite and measurable, and if one of the two functions $H(f(w)) \mid J(w) \mid$, $H(z)\kappa(z)$ is summable, then $H(z)$ belongs to C^*. Thus the assertions in §9.2 concerning the formula (32) are proved.

9.13. The formula (31) can be discussed in a similar fashion. Starting with the remark that

$$n(z) = N_{I_{+1}}(z) - N_{I_{-1}}(z)$$

almost everywhere by §5.3, we can show by a reasoning similar to that in §9.8 that (31) holds if $H(z)$ is the characteristic function of a bounded open set in the z-plane. From this we infer again (cf. the reasoning in §9.9) that (31) holds for every bounded measurable function $H(z)$.

9.14. From this it follows again (cf. §9.10) that if $H(z)$ is finite and measurable, and if $H(f(w))J(w)$, $H(z)\kappa(z)$ are both summable, then (31) holds.

9.15. From this point on, however, the argument used for (32) does not apply, since $n(z)$, $J(w)$ are not necessarily of constant sign. Assume, however, that $H(z)$ is finite and measurable, and that $H(f(w))J(w)$ is summable. Then $\mid H(f(w)) \mid \cdot \mid J(w) \mid$ is also summable, and by §9.12 this implies the summability of $\mid H(z) \mid \kappa(z)$. As $\mid n(z) \mid \leqq \kappa(z)$ almost everywhere by §6.14, it follows that $H(z)n(z)$ is also summable. Thus $H(f(w))J(w)$ and $H(z)n(z)$ are both summable, and hence (31) holds by §9.14. This completes the proof of the theorem of §9.2.

10. Special transformations

10.1. Given the continuous transformation T as in §4.1, let us assume that the partial derivatives x_u, x_v, y_u, y_v exist almost everywhere in S_0. We take an interior point (u_0, v_0) of S_0 and we put[21]

$$x(u_0, v_0) = x^0, \quad y(u_0, v_0) = y^0, \quad x_u(u_0, v_0) = x_u^0, \quad x_v(u_0, v_0) = x_v^0,$$
$$y_u(u_0, v_0) = y_u^0, \quad y_v(u_0, v_0) = y_v^0, \quad x_u^0 y_v^0 - x_v^0 y_u^0 = \Delta^0.$$

We shall consider the auxiliary transformation

$$T^*: \begin{cases} x = x^0 + (u - u_0)x_u^0 + (v - v_0)x_v^0, \\ y = y^0 + (u - u_0)y_u^0 + (v - v_0)y_v^0. \end{cases}$$

10.2. Let us put, for $(u, v) \neq (u_0, v_0)$,

$$\xi(u, v; u_0, v_0) = \frac{x(u, v) - x^0 - (u - u_0)x_u^0 - (v - v_0)x_v^0}{[(u - u_0)^2 + (v - v_0)^2]^{\frac{1}{2}}},$$

$$\eta(u, v; u_0, v_0) = \frac{y(u, v) - y^0 - (u - u_0)y_u^0 - (v - v_0)y_v^0}{[(u - u_0)^2 + (v - v_0)^2]^{\frac{1}{2}}}.$$

[21] (u_0, v_0) is a point of the set where x_u, x_v, y_u, y_v exist.

If

(37) $\xi(u, v; u_0, v_0) \to 0$, $\eta(u, v; u_0, v_0) \to 0$ for $(u, v) \to (u_0, v_0)$,

then $x(u, v)$ and $y(u, v)$ are termed *totally differentiable* at (u_0, v_0). Rademacher and subsequently a number of authors established various important geometrical results by methods based on the notion of total differentiability.[22] We shall state presently a few simple facts which played an important part in this line of work and which we shall apply then to the transformations studied in this paper.

10.3. Suppose $x(u, v)$ and $y(u, v)$ are totally differentiable at (u_0, v_0). Take a sequence of closed squares s_j such that $s_j \subset S_0$, $(u_0, v_0) \in s_j$, $|s_j| \to 0$. The following facts are then readily established by comparing T with the auxiliary transformation T^*.

(a) $G(s_j)/|s_j| \to |\Delta^0|$ (see Rademacher [1]).

(b) $\Gamma(s_j)/|s_j| \to |\Delta^0|$ (see Radó [1]).

(c) If $\Delta^0 \neq 0$, then $n_{s_j}(x^0, y^0) = \operatorname{sgn} \Delta^0$ for large j (Rademacher [1], Schauder [1]).[23]

These relations are fundamental in the study of many problems concerned with situations where $x(u, v)$, $y(u, v)$ are totally differentiable almost everywhere in S_0. According to Rademacher [1], this condition is satisfied whenever $x(u, v)$ and $y(u, v)$ satisfy the Lipschitz condition. In this case it follows from the relations (a) and (c) in §10.3 that the generalized Jacobian of §7.10 is equal to the ordinary Jacobian $\Delta = x_u y_v - x_v y_u$ almost everywhere in S_0 (Rademacher [1], Schauder [1]). We shall obtain a generalization of this result by means of the following remarks.

10.4. Let us say that a sequence $\{s_j\}$ of closed squares s_j satisfies the condition $C(\{s_j\}; u_0, v_0)$ if the following statements are true.

(i) $s_j \subset S_0$.

(ii) $(u_0, v_0) \in s_j$.

(iii) $|s_j| \to 0$.

(iv) (u_0, v_0) is not on the perimeter b_j of s_j.

(v) We have

$$\max_{(u,v)\in b_j} |\xi(u, v; u_0, v_0)| \to 0, \qquad \max_{(u,v)\in b_j} |\eta(u, v; u_0, v_0)| \to 0.$$

10.5. Inspection of the very elementary proofs[24] of the relations (a), (b), (c) in §10.3 shows that in proving (b) and (c) we are using the assumptions (37) only for points (u, v) on the *perimeters* of the squares s_j. It follows that *the*

[22] See Saks [1], Chapter 9 for references.

[23] If $\Delta^0 > 0$, then $\operatorname{sgn} \Delta^0 = +1$, and if $\Delta^0 < 0$, then $\operatorname{sgn} \Delta^0 = -1$.

[24] Cf. the detailed presentation in Radó [1].

relations (b) *and* (c) *in* §10.3 *hold whenever the sequence* $\{s_j\}$ *satisfies the condition* $C(\{s_j\}; u_0, v_0)$ *of* §10.4.

10.6. This remark suggests replacing total differentiability by a weaker condition which we shall call condition $C(T; u_0, v_0)$. We shall say that the continuous transformation T satisfies the condition $C(T; u_0, v_0)$ if (i) x_u, x_v, y_u, y_v exist at (u_0, v_0) and (ii) there exists *some* sequence $\{s_j\}$ which satisfies the condition $C(\{s_j\}; u_0, v_0)$ of §10.4.

10.7. The work of Rademacher on total differentiability was continued and generalized by Stepanoff who investigated *approximate differentiability.*[25] Stepanoff showed that *approximate differentiability, almost everywhere, is a consequence of the existence, almost everywhere, of the approximate partial derivatives of the first order.* If we assume the existence, almost everywhere, of the partial derivatives of the first order *in the usual sense,* then the reasoning which leads to the theorem of Stepanoff yields, after rather obvious modifications, the corollary that *if* x_u, x_v, y_u, y_v *exist almost everywhere in* S_0, *then the continuous transformation* T *of* §4.1 *satisfies the condition* $C(T; u_0, v_0)$ *of* §10.6 *at almost every point* (u_0, v_0) *of* S_0.

10.8. In view of the excellent presentation of the Rademacher-Stepanoff theory in Saks [1], Chapter 9, we restrict ourselves to a brief sketch of the proof. Assuming that x_u, x_v, y_u, y_v exist almost everywhere in S_0, the reasoning in Saks [1], pp. 301–303 leads to the following first result. Given $\epsilon > 0$, $\tau > 0$, there exists a closed subset E of S_0 and a $\sigma > 0$ such that the following statements are true.

(I) $|S_0 - E| < \epsilon$.

(II) We have $|\xi(u, v; u_0, v_0)| < \tau$, $|\eta(u, v; u_0, v_0)| < \tau$ for all pairs of points (u, v), (u_0, v_0) for which (a) $(u_0, v_0) \in E$, (b) $(u_0, v) \in E$, (c) $[(u - u_0)^2 + (v - v_0)^2]^{\frac{1}{2}} < \sigma$.

10.9. The preceding condition (b) is unsymmetrical. If in the reasoning of Saks, loc. cit., we exchange u and v, then it follows that in §10.8 the condition (b) can be replaced by the weaker condition (b*): *At least one* of the points (u_0, v), (u, v_0) is in E.

10.10. We shall say that a closed square s satisfies the condition $C(s; u_0, v_0; E)$ if (i) (u_0, v_0) is the center of s, (ii) $(u_0, v_0) \in E$, (iii) the sides of s are parallel to the u- and v-axes, respectively. (iv) The points of intersection of the perimeter of s with the lines $u = u_0$ and $v = v_0$ are points of E.

As almost every point of a measurable set E is a point of linear density both in the direction of the u-axis and in that of the v-axis (see Saks [1], p. 298,

[25] See Saks [1], Chapter 9.

Theorem 11.1), it follows immediately that for almost every point (u_0, v_0) of a measurable set E there exists a sequence of closed squares s_j such that $|s_j| \to 0$ and s_j satisfies the condition $C(s_j; u_0, v_0; E)$ $(j = 1, 2, \cdots)$.

10.11. Suppose s satisfies the condition $C(s; u_0, v_0; E)$ of §10.10. Clearly, for every point (u, v) on the perimeter of s at least one of the points (u_0, v), (u, v_0) is then in E. Hence the result of §10.9 can be restated as follows. Given $\epsilon > 0$, $\tau > 0$, we have a closed subset E of S_0 and a $\rho > 0$ such that

(I) $|S_0 - E| < \epsilon$, and

(II) $|\xi(u, v; u_0, v_0)| < \tau$, $|\eta(u, v; u_0, v_0)| < \tau$ for every point (u, v) on the perimeter of every square s which satisfies the condition $C(s; u_0, v_0; E)$ and also the condition $|s| < \rho$.

10.12. Take now any $\eta > 0$. Choose a sequence $\epsilon_j > 0$ such that $\epsilon_1 + \epsilon_2 + \cdots < \eta$ and a sequence $\tau_j > 0$ such that $\tau_j \to 0$. Denote by E_j the subset of S_0 which corresponds to the constants ϵ_j, τ_j in the sense of §10.11, and put $E = \prod E_j$ $(j = 1, 2, \cdots)$. Then $|S_0 - E| < \eta$, and if (u_0, v_0) is a point of E, then obviously every sequence $\{s_j\}$, such that $|s_j| \to 0$, satisfies the condition $C(\{s_j\}; u_0, v_0)$ of §10.4 as soon as each s_j satisfies the condition $C(s_j; u_0, v_0; E)$ of §10.10. By §10.10 it follows that the continuous transformation T satisfies the condition $C(T; u_0, v_0)$ of §10.6 at almost every point of E. As $|S_0 - E| < \eta$ and $\eta > 0$ was arbitrary, the statement at the end of §10.7 is proved.

10.13. Theorem. *If the continuous transformation T, given as in* §4.1, *is AC, and if the partial derivatives x_u, x_v, y_u, y_v exist almost everywhere in S_0, then the generalized Jacobian of* §7.10 *is equal to the ordinary Jacobian $\Delta = x_u y_v - x_v y_u$ almost everywhere in S_0.*

Proof. For almost every point (u_0, v_0) of S_0 the following conditions are satisfied.

(a) $\Delta(u_0, v_0)$ exists, by assumption.

(b) $J(u_0, v_0)$ exists, by §§7.10 and 7.1.

(c) If $J(u_0, v_0) \neq 0$, then $i(u_0, v_0) = \pm 1$ by §7.12, and hence $\operatorname{sgn} J(u_0, v_0) = i(u_0, v_0)$.

(d) $\Gamma'(u_0, v_0)$ exists, by §8.9.

(e) $\Gamma'(u_0, v_0) = |J(u_0, v_0)|$ by §8.9.

(f) By §10.7 there exists a sequence $\{s_j\}$ of closed squares which satisfies the condition $C(\{s_j\}; u_0, v_0)$ of §10.4. For this sequence $\{s_j\}$ the following statements are true.

(g) If $J(u_0, v_0) \neq 0$ and if (x^0, y^0) is the image of (u_0, v_0), then $i(u_0, v_0) = n_{s_j}(x^0, y^0)$ for large j, by §5.1.

(h) If $\Delta(u_0, v_0) \neq 0$, then $n_{s_j}(x^0, y^0) = \operatorname{sgn} \Delta(u_0, v_0)$ for large j, by §10.5.

The proof of the theorem is now immediate. For a point (u_0, v_0) where the above conditions (a) to (h) hold, we have $|J(u_0, v_0)| = \Gamma'(u_0, v_0) = |\Delta(u_0, v_0)|$ by §10.5 and by the conditions (d), (e), (f) above. Thus it is sufficient to show

that $J(u_0, v_0)$ and $\Delta(u_0, v_0)$ have the same sign whenever both are different from zero, and this follows directly from the conditions (c), (g), (h) above.

Bibliography

Banach, S.

[1] *Sur une classe de fonctions d'ensemble*, Fundamenta Mathematicae, vol. 6 (1924), pp. 170–188.

[2] *Sur les lignes rectifiables et les surfaces dont l'aire est finie*, ibid., vol. 7 (1925), pp. 225–236.

Burkill, J. C.

[1] *Functions of intervals*, Proc. London Math. Soc., (2), vol. 22 (1923), pp. 275–310.

[2] *The expression of area as an integral*, ibid., pp. 311–336.

Caccioppoli, R.

[1] *Transformazioni piane, superficie quadrabili, integrali di superficie*, Rendiconti del circolo mat. di Palermo, vol. 54 (1930), pp. 1–46.

Hopf, H.

[1] *Zur Topologie der Abbildungen von Mannigfaltigkeiten*, Math. Annalen, vol. 100 (1928), pp. 579–608, and vol. 102 (1929), pp. 562–623.

v. Kerékjártó, B.

[1] *Vorlesungen über Topologie*, vol. 1, Berlin, 1923.

Rademacher, H.

[1] *Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln*, parts 1 and 2, Math. Annalen, vol. 79 (1919), pp. 340–359, and vol. 81 (1920), pp. 52–63.

Radó, T.

[1] *Über das Flächenmass rektifizierbarer Flächen*, Math. Annalen, vol. 100 (1928), pp. 445–479.

[2] *On continuous transformations in the plane*, Fundamenta Mathematicae, vol. 27 (1936), pp. 201–211.

Saks, S.

[1] *Theory of the Integral*, Monografie Matematyczne, vol. 7, Warszawa, 1937.

[2] *Sur une certaine classe de fonctions d'ensemble*, Bulletin Ac. Polonaise, 1926, pp. 103–108.

Schauder, J.

[1] *Über stetige Abbildungen*, Fundamenta Mathematicae, vol. 12 (1928), pp. 47–74.

Young, W. H.

[1] *On change of the variables in a multiple integral*, Proc. Royal Soc. London, series A, vol. 96 (1920), pp. 82–91.

[2] *Integration over the area of a curve and transformation of the variables in a multiple integral*, Proc. London Math. Soc., vol. 21 (1923), pp. 161–190.

Ohio State University.

CROSS-SECTIONS OF CURVES IN 3-SPACE

By Hassler Whitney

1. **Introduction.** We consider in this paper "regular families of curves", that is, families F of non-intersecting curves such that if two curves are sufficiently close at a point, then they remain close for a "finite time"; see [3] (i.e., the third reference below), Theorem 7A. It was shown in [3] that a cross-section may always be found through an arbitrary point of the family. If the family fills a region of Euclidean 3-space E, it is natural to suspect that *any cross-section contains a cross-section which is a 2-cell*; our object here is to prove this fact. It follows (see [3], §20) that *locally, a family of curves in E is equivalent to a family of straight lines.* The proof is arranged so that a minimum of preliminary material is assumed. We use rather fully the methods in [1], [2], and the first half of [3], and a method of proof in [4].

2. **Two types of homology.** We relate the two types of homology used in [1] and [2], and give some simple properties of the second type. For a curve (= simple closed curve) J in a closed set, $J \sim 0$ was defined in [2]. For J in a general set, say $J \sim 0$ if it is ~ 0 in some bounded closed subset. Call a chain of a subdivision from some fixed sequence of simplicial subdivisions (as in [1]) a *polygonal* chain. Say two curves are equivalent in a set G if, using fixed parametrizations $f_0(\theta)$ and $f_1(\theta)$ for them, one can be deformed into the other in G, i.e., $f_t(\theta)$ $(0 \leqq t \leqq 1)$ exists and is continuous.

Lemma 1. *If J and J' are equivalent in G, then $J \sim 0$ in G if and only if $J' \sim 0$ in G.*

This is easily seen, using [2], Lemma I, if we subdivide the θ-circle and the t-segment, and consider triangles of the forms

$$f_{t_i}(\theta_j)f_{t_{i+1}}(\theta_j)f_{t_i}(\theta_{j+1}), \qquad f_{t_{i+1}}(\theta_j)f_{t_i}(\theta_{j+1})f_{t_{i+1}}(\theta_{j+1}).$$

Lemma 2. *If $J \sim 0$ in A, and A is deformed into A', leaving J fixed, then $J \sim 0$ in A'.*

For if $L \rightarrow K$, K a 1-cycle in J, we need merely consider the deformed L.

Lemma 3. *Let J be equivalent to a polygonal J' in an open set G. Then $J \sim 0$ in G if and only if J' bounds a polygonal chain in G.*

By Lemma 1, $J \sim 0$ if and only if $J' \sim 0$. Suppose J' bounds a polygonal chain. Then, using a fine enough subdivision and [2], Lemma I, we see that $J' \sim 0$. Suppose $J' \sim 0$; say $J' \sim 0$ in the closed subset A of G. Set $\epsilon =$

Received December 3, 1937; presented to the American Mathematical Society, October 29, 1932. The theorem was first proved when the author was a National Research Fellow.

$\rho(A, E - G)$, and take a simplicial subdivision of E of norm $< \epsilon$. Take δ so that a δ-chain K as in [2], Lemma I, is $\epsilon \sim 0$ in A. A simplicial approximation of this (abstract) chain gives a polygonal chain in G bounded by the polygonal J'.

LEMMA 4. *Let B be a (finite) linear graph. Let each arc α_i of B be divided into arcs of diameter $< \epsilon$, defining an ϵ-chain K_i. Let δ be the minimum distance between any two such subarcs of B without common points. Then any $(\delta, 1)$-cycle on B is $\epsilon \sim$ some linear combination of the K_i.*

The proof of [2], Lemma I, applies almost without change.

LEMMA 5. *Let Z and Z' be closed sets with only the curve J in common. If $J \sim 0$ in $Z + Z'$, then $J \sim 0$ in one of Z, Z'.*

It is sufficient to show that for every $\epsilon > 0$ there is a 1-cycle K in J as described in [2], Lemma I, which is $\epsilon \sim 0$ in one of Z, Z'. Take a sufficiently fine K in J, bounding L in $Z + Z'$. Pushing vertices near J onto J gives an $L_1 \rightarrow K$, such that $L_1 = L_2 + L_2'$, L_2 in Z, L_2' in Z'. If $L_2 \rightarrow K_2$, $L_2' \rightarrow K_2'$, then one of K_2, K_2' is $\epsilon \sim K$ in J (see [2], Lemma I; hence $K \epsilon \sim 0$ in one of Z, Z'. (Compare the proof of [2], Lemma M.)

3. **A local separation theorem.**

We shall prove two lemmas on how a portion of a cylinder separates a region. Let H denote a curved cylinder in E, given by (t, θ), $-2 \leqq t \leqq 2$, θ on a circle. Let H_{t_1,t_2} denote that part of H with $t_1 \leqq t \leqq t_2$; set $H_t = H_{t,t}$. Let H'_{t_1,t_2} and H''_{t_1,t_2} denote those parts of H_{t_1,t_2} corresponding to two complementary arcs of the θ-circle.

LEMMA 6. *Let R be a spherical region, and let H have its ends H_{-2} and H_2 outside of R. Then arbitrarily near any point q of $H \cdot R$ there are points r_1 and r_2 in different components of $R - H$.*

Choose coördinates in H so that q is on H_0. Let L be a 1-chain in $R - H_0$ linking H_0 (see [1], Theorem X^i). Say $L = L_- + L_+$, $L_- \cdot H_{-2,0} = 0$, $L_+ \cdot H_{0,2} = 0$; we may choose L_- and L_+ so that each is bounded by $a + b$, a pair of points. If a and b are in the same component of $R - H$, there is a polygonal chain M joining them. As R is spherical, we may choose $N_1 \rightarrow L_- + M$ in R; N_1 is in $E - H_{-2}$. As $H'_{-2,0}$ is a cell, we may choose $N_2 \rightarrow L_- + M$ in $E - H'_{-2,0}$ ([1], Theorem T^i). As $H_{-2} \cdot H'_{-2,0}$ is an arc, $N_1 + N_2$ cannot link it; hence there is a chain $N_3 \rightarrow L_- + M$ in $E - (H_0 + H'_{-2,0})$([1], Corollary W^i). Similarly, using $H''_{-2,0}$, we find a chain $N_4 \rightarrow L_- + M$ in $E - H_{-2,0}$. In the same manner,[1] find $N_5 \rightarrow L_+ + M$ in $E - H_{0,2}$. Adding gives $N_4 + N_5 \rightarrow L$ in $E - H_0$, a contradiction.

LEMMA 7. *Let R and R' be concentric spherical regions, $R' \subset R$. Let $H_{-1,1}$ lie in R, and let $H_{-2,-1} + H_{1,2}$ lie in $E - R'$. Then of any three points in $R' - H$, some two can be joined in $R - H$. If a cell is omitted from $H \cdot R'$, or H is replaced by $H'_{-2,2}$, then any two points can be joined.*

Take three points a, b, c in $R' - H$, and suppose no two can be joined in $R - H$.

[1] Replace H_{-2} and $H'_{-2,0}$ by $H_{-2} + H'_{-2,0}$ and $H''_{-2,0}$, respectively.

Say $L_1 \to b + c$ in R'; L_1 is in $E - (H_{-2,-1} + H_{1,2})$. Using $H'_{-1,0}$ and $H''_{-1,0}$ in R as above, etc., we find chains

$$L_{a-} \to b + c \text{ in } R - H_{-2,0}, \qquad L_{a+} \to b + c \text{ in } R - H_{0,2},$$

$$L_{b-} \to a + c \text{ in } R - H_{-2,0}, \qquad L_{b+} \to a + c \text{ in } R - H_{0,2}.$$

Let $\bar{E}$ be the sphere formed by making all of $E - R$ a single point. In $\bar{E}$, by [1], Corollary W^i, $L_{a-} + L_{a+}$ and $L_{b-} + L_{b+}$ both link H_0. Hence (Theorem X^i) their sum L_c does not. But

$$L_c = (L_{a-} + L_{b-}) + (L_{a+} + L_{b+}) = L_{c-} + L_{c+},$$

$$L_{c-} \subset R - H_{-2,0}, L_{c+} \subset R - H_{0,2}, L_{c-} \to a + b,$$

and hence, applying [1], Corollary W^i, in $\bar{E}$, $a + b$ bounds in $R - H$, a contradiction.

To prove the last two statements, we need merely note that the part of H remaining may be built up by starting with $H_{-2} + H_2$, and adding cells, each having one or two arcs in common with the preceding set; for a 1-cycle cannot link a set of arcs.

4. **The curve** J. Let S be a cross-section through p. Follow the curve of the family F through p in both directions to the first points p' and p'' on some sphere about p; joining p' to p'' by an arc outside the sphere gives a curve J^*. Take $\lambda > 0$ so that any arc[2] of $N_\lambda(p'p'')$ has just one point in S (see [3], §§9 and 16); assume $\lambda < \frac{1}{3}\rho(p, p')$ and $< \frac{1}{3}\rho(p, p'')$. The arcs of N_λ may be oriented; for any q in $S \cdot N_\lambda$, set $q_t = g'(q, t)$ ([4], §15). We may choose the function μ ([3], §3) so that $p_{-2} = p'$, $p_2 = p''$. For any q in the set of points on N_λ, let $\phi(q)$ be the point of S on the arc of N_λ through q. ϕ is continuous (compare [3], §18).

Let R_1 be a spherical region about p in N_λ. Choose a concentric R_2 so that for any q in R_2, the arc $q\phi(q)$ is in R_1. Choose a concentric R_3 so that (using a new μ if necessary) $p_{-1}pp_1 \subset R_2$ and $p_{-2}p_{-1} + p_1p_2 \subset E - \bar{R}_3$. Choose a concentric R_4 so that the last two relations hold with p and p_t replaced by any q and q_t in $S \cdot R_4$. Choose R_5 and R_6 so that $\phi(R_5) \subset R_4$, $\phi(R_6) \subset R_5$.

Take a 1-cycle A in R_6 linking J^*; we may suppose it a polygonal curve. $A' = \phi(A)$ is in R_5. We shall find a curve J in A', J not ~ 0 in $E - J^*$. Take $\eta > 0$ so small that any two points of A' within 2η of each other may be joined by an arc in A' of diameter $< \rho = \rho(A', J^*)$. Divide A into arcs A_1, $A_2, \cdots, A_n$ so that their images $\phi(A_1), \phi(A_2), \cdots$ are of diameter $< \eta$; let $x_i = \phi(A_{i-1} \cdot A_i)$. Let B_i be an arc in A' of diameter $< \rho$ joining x_i to x_{i+1}. We may replace B_i by arcs from $B_1 + \cdots + B_i$ so that $B = B'_1 + B'_2 + \cdots + B'_n$ is a linear graph (see [4], Appendix). We may define a deformation ψ

[2] In [3], §16, (2), we assume the existence of an arc q_0q_1 with a certain property; but it is easily seen that any arc, in particular $p'p''$, may be used.

carrying A along arcs of F into A', and then along short line segments into B; all this is in $E - J^*$.

Suppose that every simple closed curve of B were ~ 0 in $E - J^*$. Then there is a bounded closed set C in $E - J^*$ in which every simple closed curve of B is ~ 0; we may include in C the set of points passed over by A in the deformation ψ. As A is not ~ 0 in $E - J^*$ (Lemma 3), there is an $\epsilon > 0$ such that for every $\delta > 0$ there is a $(\delta, 1)$-cycle on A which is not $\epsilon \sim 0$ in C. We may easily choose $\zeta_1, \zeta_2, \zeta_3$ in turn so that the following is true. Take any $(\zeta_3, 1)$-cycle K_3 on A. Using ψ (see the proof of Lemma 1), we find a $(\zeta_2, 1)$-cycle K_2 on B, with $K_2 \zeta_2 \sim K_1$ in C. By Lemma 4, K_2 is $\zeta_1 \sim K_1$, a $(\zeta_1, 1)$-cycle $=$ a linear combination of the K_i as described in the lemma. Express this cycle as a sum of circuits; each lies on a simple closed curve of B, and hence is $\epsilon \sim 0$ in C. Thus any such K_3 is ~ 0 in C, a contradiction. Therefore there is a curve J in B, J not ~ 0 in $E - J^*$.

5. **The 2-cell** Q. J is in $S \cdot R_5$ and is not ~ 0 in $E - J^*$. Joining the points of J to p by line segments gives a point set D in which $J \sim 0$; applying ϕ shows that $J \sim 0$ in $S \cdot R_4$ (Lemma 2). Let P be the component of $S \cdot R_1 - J$ which contains p. Set $Q = P + J$, $P' = S \cdot R_1 - Q$. By Lemma 5, $J \sim 0$ in one of Q, $P' + J$. As J is not ~ 0 in $S - p$, $J \sim 0$ in Q.

For any point set A in S, let $H_{t_1,t_2}(A)$ denote all points q_t, q in A, $t_1 \leqq t \leqq t_2$; set $H(A) = H_{-2,2}(A)$.

We show next that J is irreducibly ~ 0 in Q. If not, then there is a point q of P so that $J \sim 0$ in $Q - q$. Let α be an arc in P joining p to q ([2], Lemma E). $D = H(\alpha)$ is a closed 2-cell. Set

$$J_1^* = (J^* - p_{-2}p_2) + H(q) + H_{-2}(\alpha) + H_2(\alpha).$$

J_1^* is a curve containing only q in P; hence $J \sim 0$ in $E - J_1^*$. Take a polygonal J_1 equivalent to J in $E - (D + J^*)$; then J_1 bounds in $E - J_1^*$ (Lemma 3). As D is a cell, J_1 bounds in $E - D$ ([1], Theorem T^i). As $D \cdot J_1^*$ is an arc, no chain can link it; hence J_1 bounds in $E - (D + J_1^*)$ ([1], Corollary W^i). Hence $J \sim 0$ in $E - (D + J_1^*)$ (Lemma 3) and hence in $E - J^*$, a contradiction.

As Q is a continuous curve (compare [2], top of p. 269), and is a cross-section through p (for the proof that J is irreducibly ~ 0 in Q shows that Q contains all points of S near p), there remains to show only that any arc γ in Q with just its ends a and b on J divides Q.

If $Q - \gamma$ is connected, let γ' be an arc in $Q - \gamma$ joining the two arcs into which a and b divide J. Let α and β be arcs of J about a and b not touching γ'. Let A and B be cells of $H(J)$ about a and b, such that $\phi(A)$ and $\phi(B)$ are subsets of α and β. By Lemma 6, there are two points in $R_4 - H(J)$ which cannot be joined by an arc in $R_1 - H(J)$. By Lemma 7, any point of $R_3 - H(J)$ can be joined to one of these by an arc in $R_2 - H(J)$; we may now write

$$R_3 - H(J) = G + G',$$

G containing p. Then two points of the same set can be joined in $R_2 - H(J)$, and two points of different sets cannot be joined in $R_1 - H(J)$.

Also, if A is omitted, G and G' may be joined; hence there is a point a' of G' arbitrarily close to a, and we may find an arc α' joining a' to A with only an end in $H(J)$. Choose β' similarly, and let γ'' join the ends of α' and β' in $R_2 - H(J)$. Projecting $\alpha' + \beta' + \gamma''$ onto S, we find an arc γ^* in $S \cdot R_1$ with its two ends on α and β. As the points of γ^* may be joined to a' through arcs of F and of $\alpha' + \beta' + \gamma''$, and a' is in G' while p is in G, no point of γ^* may be joined to p in $R_1 - H(J)$; hence γ^* has no points in P. Now γ and γ^*, together with arcs α^* in α and β^* in β, give a curve I in $S \cdot R_1$, with only $\gamma_1 = \gamma + \alpha^* + \beta^*$ in Q.

Take q within γ. By Lemma 6, there are points r_1 and r_2 arbitrarily near q which cannot be joined in $R_1 - H(I)$; $r_1' = \phi(r_1)$ and $r_2' = \phi(r_2)$ are in P. By Lemma 7, r_1' and r_2' can be joined in $R_2 - H(\gamma_1)$; applying ϕ, we find an arc δ joining them in $S \cdot R_1 - \gamma_1$. By the choice of r_1 and r_2, r_1' and r_2' cannot be joined in $Q - \gamma_1$; hence δ has points not in Q. As P is a component of $S \cdot R_1 - J$, δ has points in J. Run along δ from r_1' and r_2' to the first points s_1 and s_2 in J. As s_1 and s_2 can be joined by an arc in $J + \gamma' - (\alpha + \beta)$, r_1' and r_2' can be joined in $Q - \gamma_1$. But this is a contradiction; hence $Q - \gamma$ is not connected. All the hypotheses of Theorem I of [2] are now satisfied; hence Q is a cross-section through p which is 2-cell.

References

1. J. W. Alexander, *A proof and extension of the Jordan-Brouwer separation theorem*, Transactions of the American Mathematical Society, vol. 23 (1922), pp. 333–349.
2. H. Whitney, *A characterization of the closed 2-cell*, Transactions of the American Mathematical Society, vol. 35 (1933), pp. 261–273.
3. H. Whitney, *Regular families of curves*, Annals of Mathematics, vol. 34 (1933), pp. 244–270.
4. R. L. Wilder, *On the linking of Jordan continua in E_n by $(n-2)$-cycles*, Annals of Mathematics, vol. 34 (1933), pp. 441–449. We note a typographical error. On p. 449, line 17, $\lambda \subset t_h \cdot t_i$ should read $\lambda \supset t_h \cdot t_i$.

Harvard University.

TAUBERIAN THEOREMS FOR $(C, 1)$ SUMMABILITY

By R. P. Boas, Jr.

Let $\sum_{n=1}^{\infty} u_n$ be an infinite series, with partial sums $s_n = u_1 + u_2 + \cdots + u_n$.[1] It is well known that the applicability of Cesàro summation to the series is limited in various ways. On the one hand, the u_n cannot be too large if the series is to be summable at all;[2] on the other hand, it was shown by G. H. Hardy that if the u_n are too small, the series cannot be summable without being convergent.[3] The object of this note is to point out that in addition the Cesàro means of the series cannot approach a limit very rapidly unless the series is convergent. For simplicity, we restrict ourselves to $(C, 1)$ summability; we write $\sigma_n = n^{-1}(s_1 + s_2 + \cdots + s_n)$; then the given series is summable $(C, 1)$ to s if $\lim_{n\to\infty} \sigma_n = s$. Our theorem is

THEOREM 1. *If, as $n \to \infty$, $\sigma_n - s = o(n^{-\epsilon})$ $(0 \leqq \epsilon < 1)$ and $u_n < O(n^{\epsilon-1})$, then $s_n \to s$; if $\sigma_n - s = O(n^{-\epsilon})$ $(0 < \epsilon \leqq 1)$ and $u_n < o(n^{\epsilon-1})$, then $s_n \to s$.*

For $\epsilon = 0$, we have the known one-sided generalization of Hardy's theorem. The first part of Theorem 1 would be trivial for $\epsilon = 1$; the second part would be false for $\epsilon = 0$.

The integral analogue of Theorem 1 is

THEOREM 2. *If $g(t)$ is the derivative of its integral on every finite interval $(0, x)$, and if, as $x \to \infty$,*

$$\int_0^x (1 - x^{-1}t)g(t)\,dt - s = o(x^{-\epsilon}) \qquad (0 \leqq \epsilon < 1)$$

and $g(x) < O(x^{\epsilon-1})$; or if

$$\int_0^x (1 - x^{-1}t)g(t)\,dt - s = O(x^{-\epsilon}) \qquad (0 < \epsilon \leqq 1)$$

and $g(x) < o(x^{\epsilon-1})$; then $\int_0^{\infty} g(t)dt = s$.

The case $\epsilon = 0$ is known.[4]

Received January 12, 1938. The author is a National Research Fellow.

[1] All numbers in this note are real.

[2] If the series is summable (C, r) $(r > -1)$, $u_n = o(n^r)$ $(n \to \infty)$. See, for example, E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 2, 1926, p. 77.

[3] See E. W. Hobson, op. cit., p. 81.

[4] See E. W. Hobson, op. cit., p. 388.

We say that a function $\varphi(x)$ belongs to the class K if $\varphi(x) > 0$ $(0 \leqq x < \infty)$ and if $\varphi(cx)/\varphi(x)$ is bounded, uniformly for $0 \leqq x < \infty$ and $\frac{1}{2} \leqq c \leqq 2$. Theorem 2 is an easy consequence of the known

THEOREM 3.[5] *If $f(x)$ has a second derivative on $(0, \infty)$, and if $\varphi(x)$ and $\psi(x)$ belong to K, then, as $x \to \infty$,*

(a) *$f = O(\varphi)$, $f'' < O(\psi)$, and $\varphi = O(x^2\psi)$ imply $f' = O[(\varphi\psi)^{\frac{1}{2}}]$;*

(b) *$f = o(\varphi)$, $f'' < O(\psi)$, and $\varphi = O(x^2\psi)$ imply $f' = o[(\varphi\psi)^{\frac{1}{2}}]$;*

(c) *$f = O(\varphi)$, $f'' < o(\psi)$, and $\varphi = o(x^2\psi)$ imply $f' = o[(\varphi\psi)^{\frac{1}{2}}]$.*

We shall derive Theorem 1 from the "discontinuous" analogue of Theorem 3. This is

THEOREM 4. *If $\{a_n\}$ is a sequence of numbers, and $\varphi(n)$ and $\psi(n)$ belong to K, then, as $n \to \infty$,*

(a) *$a_n = O(\varphi)$, $\Delta^2 a_n < O(\psi)$, and $\varphi = O(n^2\psi)$ imply $\Delta a_n = O[(\varphi\psi)^{\frac{1}{2}}]$;*

(b) *$a_n = o(\varphi)$, $\Delta^2 a_n < O(\psi)$, and $\varphi = O(n^2\psi)$ imply $\Delta a_n = o[(\varphi\psi)^{\frac{1}{2}}]$;*

(c) *$a_n = O(\varphi)$, $\Delta^2 a_n < o(\psi)$, and $\varphi = o(n^2\psi)$ imply $\Delta a_n = o[(\varphi\psi)^{\frac{1}{2}}]$.*

Here $\Delta a_n = a_{n+1} - a_n$, $\Delta^2 a_n = \Delta a_{n+1} - \Delta a_n$.

If $\varphi(n)$ and $\psi(n)$ both increase, or both decrease, instead of belonging to K, and if the sign "$<$" is then replaced by "$=$" in Theorem 4, the conditions $\varphi = O(n^2\psi)$, $\varphi = o(n^2\psi)$ can be eliminated, just as in the corresponding known theorems derived in the same way from Theorem 3. From the modified theorems, as well as from the theorems as stated, more general results than Theorems 1 and 2 can easily be obtained.

To derive Theorem 2 from Theorem 3, we take in Theorem 3

$$f(x) = \int_0^x (x - t)g(t)\,dt - sx,$$

$$\varphi(x) = (1 + x)^{1-\epsilon}, \qquad \psi(x) = (1 + x)^{\epsilon-1}.$$

Then $\varphi(x)$ and $\psi(x)$ belong to K; $x^{-2}\varphi(x)/\psi(x) = O(x^{-2\epsilon}) = o(1)$ or $O(1)$ according as $\epsilon > 0$ or $\epsilon = 0$; $f'(x) = \int_0^x g(t)dt - s$; $f''(x) = g(x)$; application of (b) or (c)[6] of Theorem 3 gives

[5] R. P. Boas, Jr., *Asymptotic relations for derivatives*, this Journal, vol. 3 (1937), pp. 637–646; 638. The theorem is stated there for $f(x)$ of class C^2; inspection of the proof shows that it is unnecessary to assume that $f''(x)$ is continuous. The theorem goes back, in essentials, to G. H. Hardy and J. E. Littlewood, *Contributions to the arithmetic theory of series*, Proceedings of the London Mathematical Society, (2), vol. 11 (1912–13), pp. 411–478; 424–425; and *Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive*, Proceedings of the London Mathematical Society, (2), vol. 13 (1913–14), pp. 174–191; 188.

[6] Part (a) of Theorems 3 and 4 is not used in this note, but is included for symmetry. It can be used to establish the following companion to Theorem 1: *if $\sigma_n = O(n^{-\epsilon})$ $(0 \leqq \epsilon \leqq 1)$ and $u_n < O(n^{\epsilon-1})$, then $s_n = O(1)$.*

$$\int_0^x g(t)\,dt - s = o(1),$$

which is the conclusion of Theorem 2.

Similarly, to derive Theorem 1 from Theorem 4, we take in Theorem 4

$$a_n = n\sigma_n - ns,$$
$$\varphi(n) = (1 + n)^{1-\epsilon}, \qquad \psi(n) = (1 + n)^{\epsilon-1}.$$

Then $\Delta a_n = s_n - s$, $\Delta^2 a_n = u_n$, and Theorem 1 follows.

We now establish Theorem 4. In order to treat the three cases simultaneously, we suppose, as we clearly may, that

$$(1)\qquad \begin{aligned} \Delta^2 a_n &< \lambda(n)\psi(n), \\ |\, a_n \,| &< \mu(n)\varphi(n), \end{aligned}$$

where $\lambda(n)$ and $\mu(n)$ are non-increasing.

We take $n > 4$ and $k < \frac{1}{2}n$. Since

$$(2)\qquad \Delta a_{n+1} - \Delta a_n = \Delta^2 a_n < \lambda(n)\psi(n),$$

we have, replacing n successively by $n - 1, n - 2, \cdots, n - k + 1$, and using (2) repeatedly,

$$\Delta a_n < \Delta a_{n-1} + \lambda(n - 1)\psi(n - 1),$$
$$\Delta a_n < \Delta a_{n-2} + \lambda(n - 2)\psi(n - 2) + \lambda(n - 1)\psi(n - 1),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\Delta a_n < \Delta a_{n-k} + \lambda(n - k)\psi(n - k) + \lambda(n - k + 1)\psi(n - k + 1) + \cdots + \lambda(n - 1)\psi(n - 1).$$

Adding these inequalities, we obtain, since $\Delta a_n = a_{n+1} - a_n$,

$$(3)\qquad k\Delta a_n < a_n - a_{n-k} + \sum_{j=1}^{k} (k - j + 1)\lambda(n - j)\psi(n - j).$$

Similarly, using (2) with n replaced successively by $n + 1, n + 2, \cdots, n + k$, we obtain

$$\Delta a_n > \Delta a_{n+1} - \lambda(n)\psi(n),$$
$$\Delta a_n > \Delta a_{n+2} - \lambda(n + 1)\psi(n + 1) - \lambda(n)\psi(n),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\Delta a_n > \Delta a_{n+k} - \lambda(n + k - 1)\psi(n + k - 1) - \lambda(n + k)\psi(n + k) - \cdots - \lambda(n)\psi(n).$$

Adding, we have

$$(4)\qquad k\Delta a_n > a_{n+k+1} - a_{n+1} - \sum_{j=0}^{k-1} (k - j)\lambda(n + i)\psi(n + j).$$

Now, $| a_n | < \mu(n)\varphi(n)$, and, for some A, and for $\frac{1}{2}n < k < n$,

$$\frac{\psi(n \pm j)}{\psi(n)} < A, \qquad \frac{\varphi(n \pm j)}{\varphi(n)} < A \quad (j = 0, 1, 2, \cdots, k).$$

Therefore (3) and (4) yield

$$k\Delta a_n < A\varphi(n)[\mu(n) + \mu(n - k)] + A\psi(n) \sum_{j=1}^{k} (k - j + 1)\lambda(n - j),$$

$$k\Delta a_n > -A\varphi(n)[\mu(n + k + 1) + \mu(n + 1)] - A\psi(n) \sum_{j=0}^{k-1} (k - j)\lambda(n + j).$$

Since $\lambda(n)$ and $\mu(n)$ are non-increasing, we obtain, combining the last two inequalities,

$$k\,| \Delta a_n | < 2A\varphi(n)\mu(\tfrac{1}{2}n) + A\psi(n)\lambda(\tfrac{1}{2}n) \sum_{j=0}^{k-1} (k - j),$$

$$(5) \qquad | \Delta a_n | < 2Ak^{-1}\varphi(n)\mu(\tfrac{1}{2}n) + Ak\psi(n)\lambda(\tfrac{1}{2}n).$$

If we take $k = k(n) = \omega(n)[\varphi(n)/\psi(n)]^{\frac{1}{2}}$, where $\omega(n)$ is a positive function such that $k(n) < \frac{1}{2}n$, (5) becomes

$$(6) \qquad | \Delta a_n |\, [\varphi(n)\psi(n)]^{-\frac{1}{2}} < 2A\mu(\tfrac{1}{2}n)\omega(n)^{-1} + A\omega(n)\lambda(\tfrac{1}{2}n).$$

For (a), we may suppose $\lambda(n)$ and $\mu(n)$ to be constants, and take for $\omega(n)$ a bounded function, bounded from zero, such that $k(n)$ is an integer less than $\frac{1}{2}n$; then (6) gives the conclusion.

For (b), $\mu(n) \to 0$, and we may suppose $\lambda(n)$ constant. If $\omega(n)$ is chosen not greater than $\mu(\frac{1}{2}n)^{\frac{1}{2}}$ and so that $k(n) < \frac{1}{2}n$, the right side of (6) approaches zero $(n \to \infty)$.

For (c), we have $[\varphi(n)/\psi(n)]^{\frac{1}{2}} = \alpha(n) = o(n)$; $\mu(n)$ a constant; $\lambda(n) = o(1)$. It is then possible to choose $\omega(n)$ so that $\omega(n) \to \infty$, while $\omega(n)n^{-1}\alpha(n) \to 0$ and $\omega(n)\lambda(\frac{1}{2}n)^{\frac{1}{2}} \to 0$. Then $k(n) < \frac{1}{2}n$ for large n, and the right side of (6) approaches zero.

PRINCETON UNIVERSITY.

THE INDEX THEOREM IN THE CALCULUS OF VARIATIONS

By Marston Morse

Introduction. The calculus of variations in the large is concerned with boundary problems in the large. Of these problems the simplest is that of finding extremals joining two points A and B on a regular m-manifold M. The theory[1] obtains relations between the local characteristics of the solutions of the problem and the topological characteristics (connectivities, etc.) of the space of admissible curves. The case where A and B are not conjugate on any extremal solution is termed the *non-degenerate case*. This case is the general case in the sense that for A fixed the set of points B which are conjugate to A on at least one extremal issuing from A has a null m-dimensional measure on M. The unrestricted case can be treated as a limiting case of the non-degenerate case (M, p. 239).

It appears that the most significant characteristic of an extremal solution g in the non-degenerate case is the number μ of conjugate points of A on g. This fact becomes most evident in terms of the "Index Theorem". Recall that the "index" of a critical point $(z) = (0)$ of a function $J(z)$ of a finite number of variables (z) is the number of negative characteristic roots of the Hessian of $J(z)$ at the point $(z) = (0)$. As we shall see $J(z)$ will represent the value of the integral J along a "canonical" broken extremal neighboring g with vertices determined by (z). The Index Theorem affirms that μ equals the index of the critical point (0) of this function $J(z)$. It is by means of this theorem that the topological characteristics of the neighborhood of g among admissible curves are determined.

The Index Theorem was first established in the non-parametric case in 1929 by Morse.[2] It was established in the parametric case by a reduction to the non-parametric case (M, p. 138). Recently the author has discovered a new and simpler method of proving the theorem. This method can be applied

Received December 22, 1937.

[1] M. Morse, *The Calculus of Variations in the Large*, American Mathematical Society Colloquium Publications, vol. 18, New York, 1934. A reference to these lectures will be indicated by the letter M.

Classical treatments of the conjugate point condition can be found in the following references.

G. A. Bliss, *Jacobi's condition for problems of the calculus of variations in parametric form*, Transactions of the American Mathematical Society, vol. 17 (1916), pp. 195–206.

C. Carathéodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, Berlin, Teubner, 1935.

[2] M. Morse, *The foundations of the calculus of variations in the large in m-space*, Transactions of the American Mathematical Society, vol. 31 (1929), pp. 379–404.

directly to the parametric case and will be presented in this paper. The lemmas involved moreover have been useful in making advances in the theory of the "index form" under general boundary conditions.

1. **The space and integral.** We shall be concerned with a regular m-manifold M of class C^{r+1} with $r > 2$. Such a manifold is a Hausdorff topological space with the following properties. If p is a point of M, there exists a neighborhood of p which is the homeomorph of a region N in a Euclidean space E of dimension m, with coördinates (x) (termed *preferred* coördinates), such that any two sets (x) and (z) of preferred coördinates neighboring a point p on M are related by a non-singular transformation

$$z^i = z^i(x) \qquad (i = 1, \cdots, m) \tag{1.1}$$

of class C^{r+1}. Any system of coördinates obtained from preferred coördinates (x) by a non-singular transformation of the form (1.1) and of class C^n, $0 < n \leqq r + 1$, will be termed a C^n coördinate system.

In each preferred system (x) we suppose there is defined a function

$$F(x^1, \cdots, x^m, r^1, \cdots, r^m) = F(x, r)$$

of class C^r in (x, r) for (x) in the system (x) and for all sets $(r) \neq (0)$. We suppose that F is an invariant. More precisely, if $Q(z, \sigma)$ replaces F in a preferred system (z) and (x) and (z) are related as in (1.1), and if (r) in the system (x) is transformed as a contravariant tensor into (σ) in the system (z), then

$$F(x, r) = Q(z, \sigma).$$

We suppose that $F(x, r)$ is positive homogeneous of order 1 in the variables (r). Recall that

$$| F_{r^i r^j}(x, r) | = 0 \qquad (i, j = 1, \cdots, m). \tag{1.2}$$

We assume that the rank of the determinant (1.2) is $m - 1$. (Cf. M, pp. 111–112).

A subset of M locally representable in the form

$$x^i = x^i(u_1, \cdots, u_p) \qquad (i = 1, \cdots, m)$$

in terms of functions of class C^n, $n > 0$, with a functional matrix of rank p will be termed a *regular* m-manifold of class C^n on M. A regular 1-manifold (or curve) on M of class C^2 with a local parameter t will be called an *extremal* if it satisfies the Euler equations

$$\frac{d}{dt} F_{r^i} - F_{x^i} = 0 \qquad (i = 1, \cdots, m) \tag{1.3}$$

in each coördinate system (x) in which it enters. Such an extremal can be shown to be a regular curve of class C^r in terms of a suitably chosen parameter. We shall suppose our extremals so represented.

Let g be a closed arc of such an extremal. If $\gamma^i(t)$ is a regular representation of g in the system (x), then

$$F^0_{r^i r^j} \dot{\gamma}^j(t) \equiv 0, \tag{1.4}$$

where the superscript 0 indicates that the arguments are $\gamma^i(t)$ and $\dot{\gamma}^i(t)$. We assume that

$$F^0_{r^i r^j} w^i w^j > 0 \tag{1.5}$$

for every non-null set (w) independent of $[\dot{\gamma}(t)]$.

Recall that any extremal arc g lies in a C^r coördinate system (x). (Cf. M, p. 108, Theorem 1.1.) We shall define the conjugate points of a point A on g. Suppose that $t = t_0$ at A. The extremals issuing from A with directions neighboring that of g can be represented in the system (x) in the form (M, p. 117),

$$x^i = x^i(t, u) \qquad (i = 1, \cdots, m), \tag{1.6}$$

where (u) is a set of $n = m - 1$ parameters u_j and is constant on each extremal, and where t is a parameter in terms of which the extremal (u) is regularly represented. We suppose that the set $(u) = (0)$ determines g and that on g, $t_1 \leqq t \leqq t_2$. A representation of this character exists in which the functions $x^i(t, u)$ and $x^i_t(t, u)$ are of class C^{r-1} in terms of their arguments for (u) neighboring $(u) = (0)$ and t on any interval $(t_1 - e, t_2 + e)$ for which e is positive and sufficiently small, while the Jacobian

$$\Delta(t, t_0) = \frac{D(x^1, \cdots, x^m)}{D(t, u_1, \cdots, u_n)} \not\equiv 0, \qquad (u) = (0), \tag{1.7}$$

on g neighboring $t = t_0$. In terms of any representation of this character we define the conjugate points of $t = t_0$ on g as the points t on g at which $\Delta(t, t_0) = 0$. The order of vanishing of $\Delta(t, t_0)$ at a conjugate point t will be termed the *order* of that conjugate point. It is easy to show that the conjugate points and their orders are independent of the particular coördinate system (x) in which g lies and of representations of the character (1.6) in terms of which the conjugate points are defined.

2. **The second variation.** We suppose that $F(x, \dot{x}) > 0$ along g. No generality is lost in making this assumption. If this hypothesis were not satisfied by a given integrand $G(x, \dot{x})$, we could set

$$F(x, \dot{x}) = G(x, \dot{x}) + k^2 \dot{x}^i \dot{\gamma}^i(t),$$

and obtain a new integrand F which would be positive along g provided k were sufficiently large. Moreover the extremals and conjugate points corresponding to the two integrands would be the same, and F and G would satisfy the condition (1.5) together.

With $F > 0$ along g we can represent any curve h on which (x) and $(\dot{x})$ are sufficiently near similar elements on g in terms of a parameter t which equals

the value of J along h. Such a parameter t will be called the J-length along h. Along a curve on which t is the *J-length* F will be identically 1. We suppose g represented in terms of its J-length in the form

$$x^i = \gamma^i(t) \qquad (t_1 \leqq t \leqq t_2). \tag{2.0}$$

Recall that the second variation in the fixed end point problem corresponding to an extremal arc g has the form

$$I(\eta) = \int_{t_1}^{t_2} 2\Omega(\eta, \dot\eta)\,dt, \tag{2.1}$$

where

$$2\,\Omega(\eta, \dot\eta) = F^0_{x^ix^j}\eta^i\eta^j + 2F^0_{x^ir^j}\eta^i\dot\eta^j + F^0_{r^ir^j}\dot\eta^i\dot\eta^j, \tag{2.2}$$

and that the Jacobi equations (written J. E.) have the form

$$L_i(\eta) = \frac{d}{dt}\Omega_{\dot\eta^i} - \Omega_{\eta^i} = 0 \qquad (i = 1, \cdots, m). \tag{2.3}$$

As is well known any two solutions $u^i(t)$ and $z^i(t)$ of the J. E. of class C^2 satisfy the integral

$$u^i\Omega_{\dot\eta^i}(z, \dot z) - z^i\Omega_{\dot\eta^i}(u, \dot u) \equiv \text{constant}. \tag{2.4}$$

Recall also that the equations $L_i(\eta) = 0$ are not independent, in fact satisfy the relation (cf. M, p. 123)

$$\dot\gamma^i(t)L_i(\eta) \equiv 0 \tag{2.5}$$

for all functions $\eta^i(t)$ of class C^2. Because of this we replace the system (2.3) by the system

$$L_i(\eta) = 0, \qquad \frac{d}{dt}M(\eta) = 0, \qquad (i = 1, \cdots, m), \tag{2.6}$$

where

$$M(\eta) = F^0_{x^i}\eta^i + F^0_{r^i}\dot\eta^i. \tag{2.7}$$

The operator $M(\eta)$ is suggested as follows. If $x^i(t, e)$ is a family of extremals of class C^2 yielding g for $e = 0$, we have

$$\frac{\partial}{\partial e}F[x(t, e), \dot x(t, e)]\Big|^0 = M(\eta), \tag{2.8}$$

where

$$\eta^i = x^i_e(t, 0). \tag{2.9}$$

In particular if $F \equiv 1$ on members of the family, the corresponding variations η^i will satisfy the system (2.6). We term the system (2.6) the *restricted* Jacobi equations (written R. J. E.).

The R. J. E. are equivalent to the system

$$(2.10)' \qquad L_i(\eta) + \mu F^0_{r^i} = 0,$$

$$(2.10)'' \qquad \frac{d}{dt} M(\eta) = 0,$$

as a set of conditions on (η). For if we multiply the i-th equation in $(2.10)'$ by $\dot\gamma^i$ and sum, it follows from (2.5) and the relation

$$F^0_{r^i}\dot\gamma^i \equiv F^0 \equiv 1$$

that $\mu \equiv 0$. The functional determinant of the system (2.10) with respect to the variables $\ddot\eta^i$ and μ is

$$\begin{vmatrix} F_{r^i r^j} & F_{r^i} \\ F_{r^j} & 0 \end{vmatrix}^0 = -F^0_1(F^0_{r^i}\dot\gamma^i)^2 = -F^0_1 \neq 0,$$

in accordance with M, p. 112. Hence the R. J. E. can be written in the form

$$(2.11) \qquad \ddot\eta^i = H^i(\eta, \dot\eta),$$

where the functions H^i are linear and homogeneous in the variables η^i, $\dot\eta^i$ with coefficients which are functions of t of class C^{r-2}.

If $\rho(t)$ is any function of class C^2,

$$(2.12) \qquad \eta^i = \rho(t)\dot\gamma^i(t)$$

is a solution of the J. E., (2.3). In fact for values of the constant e sufficiently near zero the functions

$$(2.13) \qquad x^i = \gamma^i[t + e\rho(t)]$$

represent g and hence satisfy the Euler equations. According to the theorem of Jacobi the partial derivatives of the right members of (2.13) with respect to e, evaluated for $e = 0$, afford a solution of the J. E. These partial derivatives reduce to the form (2.12), and our statement is proved. We term solutions of the J. E. of the form (2.12) *tangential variations.* We continue with the following lemma.

LEMMA 2.1. *The only tangential variations which are solutions of the R. J. E. are of the form* $(a + bt)\dot\gamma^i$ *where a and b are constants.*

For functions $\eta^i(t)$ of the form (2.12), $M(\eta)$ takes the form

$$(2.14) \qquad M(\eta) = F^0_{r^i}\dot\rho\dot\gamma^i + \rho[F^0_{r^i}\ddot\gamma^i + F^0_{x^i}\dot\gamma^i].$$

But the factor in the bracket vanishes, as we see upon differentiating the identity

$$F[\gamma(t), \dot\gamma(t)] \equiv 1$$

with respect to t. Hence (2.14) takes the form

$$M(\eta) \equiv \dot\rho,$$

from which the lemma follows.

We return to the determinant $\Delta(t, t_0)$ of (1.7). We suppose now that the extremals of the family (1.6) have been so represented that t is the J-length on each extremal. We then multiply the first column of $\Delta(t, t_0)$ by $t - t_0$ and denote the resulting determinant by $D(t, t_0)$. We term D the *conjugate point determinant*. Each of its columns is a solution of the R. J. E. and vanishes at $t = t_0$. The columns of D are independent solutions of the R. J. E. since $D \not\equiv 0$ for t near t_0. It follows from the form (2.11) of the R. J. E. that the columns of D form a base for solutions of the R. J. E. which vanish at $t = t_0$.

We continue with the following theorem.

THEOREM 2.1. *A necessary and sufficient condition that a point $t = a$ be conjugate to $t = t_0$ on g is that there exist a solution of the R. J. E. which vanishes at $t = t_0$ and $t = a$ but which does not vanish identically.*

To prove the condition necessary we suppose that $D(a, t_0) = 0$ with $a \neq t_0$. There will then exist a proper linear combination (η) of the columns of $D(t, t_0)$ which vanishes at $t = a$. We note that $(\eta) = (0)$ at t_0 but that $(\eta) \not\equiv (0)$ since $D(t, t_0) \not\equiv 0$. Hence the condition is necessary.

To prove the condition sufficient we assume that (w) is a solution of the R. J. E. which vanishes at $t = t_0$ and $t = a$ with $(w) \not\equiv (0)$. Since the columns of $D(t, t_0)$ form a base for solutions of the R. J. E. which vanish at $t = t_0$, (w) is dependent on the columns of D. But $(w) = (0)$ at a so that $D(a, t_0) = 0$, and the condition is proved sufficient.

Understanding that the nullity ν of a determinant equals its order minus its rank r we shall prove the following theorem. (Cf. M, p. 47, Theorem 3.1.)

THEOREM 2.2. *If $\theta(t)$ is a determinant whose columns are m independent solutions of the R. J. E. which vanish at $t = t_0$, the order of vanishing of $\theta(t)$ at any point $t = a$ equals the nullity ν of $\theta(a)$.*

Let r be the rank of $\theta(a)$ so that $\nu + r = m$. Without loss of generality we can suppose that the rank of the first r columns of $\theta(a)$ is r, for this would result after a suitable reordering of the columns. We lose no generality if we also suppose that the rank of the last ν columns of $\theta(a)$ is zero, since this would result upon adding suitable linear combinations of the first r columns to the remaining columns. With this understood let

$$u_h^i(t) \qquad (h = 1, \cdots, r),$$

$$z_k^i(t) \qquad (k = 1, \cdots, \nu)$$

represent the first r and last ν columns respectively of $\theta(t)$. Upon applying the integral form of the law of the mean to the elements in the last ν columns of $\theta(t)$, we find that

$$\theta(t) = (t - a)^\nu B(t), \tag{2.15}$$

where $B(t)$ is continuous in t and

$$B(a) = | u_h^i(a) \quad \dot{z}_k^i(a) | \qquad (h = 1, \cdots, r;\ k = 1, \cdots, \nu). \tag{2.16}$$

The theorem will follow from (2.15) if we show that $B(a) \neq 0$.

Suppose that $B(a) = 0$. I say first that $r > 0$. Otherwise there would exist a proper linear combination $\eta^i(t)$ of the columns of $\theta(t)$ such that $\eta^i(a) = \dot{\eta}^i(a) = 0$ for each i. It would follow that $(\eta) \equiv (0)$ contrary to the hypothesis that the columns of $\theta(t)$ are independent.

We suppose then that $B(a) = 0$ and $r > 0$. There will exist a proper linear combination (w) of the columns of $B(a)$ with coefficients

$$c_1, \cdots, c_r, -d_1, \cdots, -d_\nu,$$

such that $(w) = (0)$. Moreover, the constants d_k are not all null. Otherwise

$$c_h u_h^i(a) = 0 \qquad (h = 1, \cdots, r)$$

for each i, and the rank of $\theta(a)$ would be less than r.

We set

$$u^i(t) = c_h u_h^i(t), \qquad z^i(t) = d_k z_k^i(t), \qquad (i = 1, \cdots, m).$$

Then

$$u^i(a) = \dot{z}^i(a), \qquad z^i(a) = 0, \qquad (i = 1, \cdots, m). \tag{2.17}$$

We note that

$$[u(a)] \neq [0]. \tag{2.18}$$

Otherwise the relations (2.17) would imply that $[z(t)] \equiv [0]$, contrary to the hypothesis that the columns of $\theta(t)$ are independent.

The relation (2.4) is satisfied by the solutions $u^i(t)$ and $z^i(t)$ with its right member null. For $u^i(t)$ and $z^i(t)$ vanish at $t = t_0$ for each i. Recalling that

$$\Omega_{\dot{\eta}^i}(z, \dot{z}) = F^0_{r^i r^j} \dot{z}^j + F^0_{r^i z^j} z^j$$

and making use of (2.17), we see that relation (2.4) reduces to the relation

$$F^0_{r^i r^j} u^i(a) u^j(a) = 0$$

at $t = a$. It follows from (1.5) that

$$u^i(a) = c\dot{\gamma}^i(a) \qquad (c \neq 0;\ i = 1, \cdots, m).$$

This is impossible. For there would then exist a solution of the R. J. E. of the form

$$\eta^i(t) = z^i(t) - c(t - a)\dot{\gamma}^i \tag{2.19}$$

with $\eta^i(t) \equiv 0$ for each i, since

$$\eta^i(a) = \dot{\eta}^i(a) = 0 \qquad (i = 1, \cdots, m).$$

But upon setting $t = t_0$ in (2.19) we find that

$$0 = c(t_0 - a)\dot{\gamma}^i(t_0) \qquad (i = 1, \cdots, m). \tag{2.20}$$

Now $t_0 \neq a$ in the case $r > 0$, so that relations (2.20) are impossible, and the proof is complete.

The conjugate point determinant $D(t, t_0)$ satisfies the conditions on $\theta(t)$. In particular its columns are independent since $D(t, t_0) \not\equiv 0$ near $t = t_0$. We accordingly have the following corollary of the theorem.

COROLLARY 2.1. *The conjugate points of $t = t_0$ on g are isolated and possess orders equal to the nullity of the conjugate point determinant $D(t, t_0)$ at the respective zeros $t \neq t_0$ of this determinant.*

The columns of $D(t, t_0)$ form a base for solutions of the R. J. E. which vanish at $t = t_0$. The maximum number of independent solutions of the R. J. E. which vanish at $t = t_0$ and at a conjugate point $t = a$ of $t = t_0$ will accordingly equal the nullity of $D(a, t_0)$. We thus have a second corollary of the theorem.

COROLLARY 2.2. *The maximum number of independent solutions of the R. J. E. which vanish at $t = t_0$ and at a conjugate point $t = a$ of $t = t_0$ equals the order of the conjugate point $t = a$.*

The conjugate point determinant can be given a representation of the form

$$D(t, t_0) = (t - t_0)^m B(t, t_0),$$

where $B(t, t_0)$ is continuous in its arguments for t and t_0 on the interval (t_1, t_2). Moreover,

$$B(t_0, t_0) \neq 0,$$

as we have seen in connection with (2.15). It follows that the first conjugate point of $t = t_0$ following t_0 is bounded away from t_0 on g. We shall make use of this fact in the next section.

3. The Index Theorem. We continue with the extremal g represented in terms of its J-length as in (2.0). Let A and B be respectively the initial and final points of g. Let ω be a positive lower bound of J-lengths on g between a point on g and the first following conjugate point. Let

$$a_0 < a_1 < \cdots < a_{p+1} \qquad (a_0 = t_1;\ a_{p+1} = t_2)$$

be a set of values of t such that

$$a_s - a_{s-1} < \omega \qquad (s = 1, \cdots, p + 1). \tag{3.0}$$

Let $A_0, \cdots, A_{p+1}$ be the corresponding points on g. Let M_q be a regular manifold of class C^{r-1} (cf. §1) which intersects g at A_q, $q = 1, \cdots, p$, but which is not tangent to g at A_q. We suppose M_q regularly represented neighboring A_q in terms of parameters

$$u_q^j \qquad (j = 1, \cdots, n = m - 1) \tag{3.1}$$

in such a manner that the set $(u_q) = (0)$ determines A_q on g. We shall term the manifolds M_q a set of *intermediate* manifolds.

It will be convenient to write the ensemble of the sets u_q^j in the form

$$(z^1, \cdots, z^{pn}) = (u_1^1, \cdots, u_1^n, \cdots, u_p^1, \cdots, u_p^n). \tag{3.2}$$

The ensemble (z) determines a set of points

$$P_1, \cdots, P_p \tag{3.3}$$

on the respective manifolds M_q provided (z) is sufficiently near the set $(z) = (0)$.

For (z) sufficiently near (0) we can join the successive points

$$A, P_1, \cdots, P_p, B \tag{3.3$'$}$$

by extremal arcs forming a broken extremal $E(z)$. The J-length of $E(z)$ will be denoted by $J(z)$ and will be a function of class C^{r-1} of its arguments for (z) sufficiently near (0). The Index Theorem can now be stated in full as follows.

INDEX THEOREM. *The point $(z) = (0)$ is a critical point of $J(z)$ with an index equal to the number of conjugate points of $t = t_1$ on g preceding $t = t_2$, and a nullity equal to the order of $t = t_2$ as a conjugate point of $t = t_1$ on g.*

It is understood that each conjugate point is to be counted a number of times equal to its order. Recall also that the *nullity* of a critical point is defined as the nullity of the Hessian of the function at the critical point.

We shall represent the broken extremal $E(z)$ in terms of a parameter t which equals a_0 and a_{p+1} respectively at the end points of g and which equals a_q on the respective manifolds M_q. We suppose t chosen so that between any two successive vertices (3.3)$'$ the rate of change of t with respect to the J-length is constant. We thereby obtain a representation of $E(z)$ of the form

$$x^i = x^i(t, z) \qquad (t_1 \leqq t \leqq t_2), \tag{3.4}$$

where $x^i(t, z)$ is continuous in its arguments for (z) sufficiently near (0) and of class C^{r-1} for t on the respective intervals (a_k, a_{k+1}). For t fixed the functions $x^i(t, z)$ are of class C^{r-1} in (z) without exceptional values of t.

A computation involving the usual differentiation and integration by parts discloses the fact that for $(z) = (0)$,

$$\frac{\partial J}{\partial z^\mu} = \sum_{k=0}^{p} \left[F^0_{r^i} \frac{\partial x^i}{\partial z^\mu} \right]_{a_k}^{a_{k+1}} \qquad (\mu = 1, \cdots, pn). \tag{3.5}$$

Since the contributions to the right members from the fixed end points are null and the contributions from the two extremal arcs with end points on M_q cancel, it appears that $(z) = (0)$ is a critical point of $J(z)$, as stated.

We set

$$Q(z) = \frac{\partial^2 J^0}{\partial z^\mu \partial z^\nu} z^\mu z^\nu \qquad (\mu, \nu = 1, \cdots, pn), \tag{3.6}$$

where the superscript 0 indicates evaluation for $(z) = (0)$, and term $Q(z)$ an *index form* corresponding to g.

Let e be a parameter. For (z) fixed,

$$x^i = x^i(t, ez) \qquad (i = 1, \cdots, m) \tag{3.7}$$

is a 1-parameter family of broken extremals. We find that

$$\frac{d^2}{de^2} J(ez) \Big|^{e=0} = Q(z),$$

and infer the following.

THEOREM 3.1. *The index form $Q(z)$ admits the representation*

$$Q(z) = \int_{t_1}^{t_2} 2\,\Omega(\eta, \dot{\eta})\, dt, \tag{3.8}$$

with

$$\eta^i(t) = a^i_\mu(t) z^\mu \qquad (\mu = 1, \cdots, pn), \tag{3.9}$$

where

$$a^i_\mu(t) = \frac{\partial x^i(t, z)}{\partial z^\mu}\Big|^{(z)=(0)}. \tag{3.10}$$

In order to make full use of this theorem we shall need certain properties of quadratic forms. To that end let

$$q(x) = a_{ij} x_i x_j \qquad (i, j = 1, \cdots, m)$$

be a symmetric quadratic form. Let r be the rank of $| a_{ij} |$. The number $m - r$ is termed the *nullity* of q and will be denoted by $N(q)$. The *index* of $q(x)$ is the number of negative characteristic roots belonging to $| a_{ij} |$ and will be denoted by $I(q)$. Recall that the characteristic roots will vary continuously with the coefficients a_{ij}. The index $I(q)$ is an integer, and, although not a continuous function of the coefficients a_{ij}, is easily seen to be *lower semi-continuous.*

Let π_h be an h-plane passing through the origin in the space (x). The h-plane π_h can be represented in the form

$$x_i = b_{ij} u_j \qquad (j = 1, \cdots, h;\ i = 1, \cdots, m), \tag{3.11}$$

where the matrix $\| b_{ij} \|$ has the rank h. Upon replacing the variables (x) in $q(x)$ by the right members of (3.11), we obtain a form $H(u)$ in the variables (u). We term $H(u)$ a form defined by $q(x)$ on π_h. We say that $q(x)$ is negative definite or semi-definite on π_h if $H(u)$ is negative definite or semi-definite. With this understood we state the following lemma (M, p. 61, Lemma 7.1).

LEMMA 3.1. (a) *A necessary and sufficient condition that the index of $q(x)$ be at least h is that $q(x)$ be negative definite on some h-plane π_h through the origin.* (b) *A necessary and sufficient condition that the index plus the nullity of $q(x)$ be at least p is that $q(x)$ be negative semi-definite on some p-plane through the origin.*

Let $A(y_1, \cdots, y_n)$ be a quadratic form in the n variables (y) defined by the relation

$$A(y) = q(y_1, \cdots, y_n, 0, \cdots, 0) \qquad (0 < n \leqq m).$$

Concerning $A(y)$ we state the following lemma.

LEMMA 3.2. (a) *The index of the form $q(x)$ is at least that of $A(y)$.* (b) *If $q(x)$ is non-degenerate,*

$$I(q) \geqq I(A) + N(A). \tag{3.12}$$

The first statement in the lemma is a consequence of the first statement in Lemma 3.1 since $A(y)$ is negative definite on some h-plane through the origin with $h = I(A)$. To prove the relation (3.12) let p equal the right member of (3.12). According to Lemma 3.1 (b) there exists a p-plane through the origin on which $A(y)$ is negative semi-definite. Upon identifying the space (y) with the subspace of the first n coördinates x_i, we see that $q(x)$ is negative semi-definite on this p-plane. It follows from Lemma 3.1 (b) that

$$I(q) + N(q) \geqq p.$$

But $N(q) = 0$ by hypothesis, and relation (3.12) follows.

4. **Proof of the Index Theorem.** The second variation is evaluated in Theorem 3.1 along a family of curves given by (3.9). We regard the sets (z) as parameters in this family. On each interval

$$a_k \leqq t \leqq a_{k+1} \qquad (k = 0, \cdots, p), \tag{4.1}$$

it is seen that (3.9) defines a solution of the R. J. E. That the J. E. are satisfied follows from the fact that the functions $x^i(t, z)$ of (3.4) define extremal arcs λ for t on (4.1). But on each such extremal arc λ, F is identically constant by virtue of our choice of t. Hence the last condition in (2.6) is satisfied by the subarcs of (3.9) for which (4.1) holds. For (z) fixed, (3.9) thus defines a "broken" solution of the R. J. E.

For q fixed the variables u_q^j in the set (z) are parameters in our representation of M_q. Let the partial derivative of $x^i(t, z)$ with respect to u_q^j be indicated by adding the subscripts j and q. Note that the matrix (q fixed)

$$\| x_{jq}^i(a_q, 0) \| \qquad (i = 1, \cdots, m; j = 1, \cdots, n) \tag{4.2}$$

has the rank n since M_q is regularly represented in terms of the parameters u_q^j. From the fact that M_q is not tangent to g at its intersection with g it follows that

$$| \dot{\gamma}^i(a_q) \quad x_{jq}^i(a_q, 0) | \neq 0, \tag{4.3}$$

where the determinant in (4.3) is obtained from the matrix (4.2) by adjoining the column $\dot{\gamma}^i(a_q)$.

We continue with the following lemma.

LEMMA 4.1. *Linearly independent sets (z) determine linearly independent broken solutions of the R. J. E. in* (3.9).

If $(z) \neq (0)$, there will be some q for which the set u_q^j is not null. At the point $t = a_q$ on the broken solution $\eta^i(t)$ determined by (z) in (3.9) the set

$$\eta^i(a_q) = x_{jq}^i(a_q, 0)u_q^j \qquad (i = 1, \cdots, m; j = 1, \cdots, n)$$

is not null since the rank of the matrix (4.2) is n. The lemma follows directly.

We shall say that a broken solution of the family (3.9) is a *special* solution if corresponding to each corner $t = a_q$ there is a constant c such that

$$\Delta\dot{\eta}^i = \dot{\eta}^i \Big|_{a_q^-}^{a_q^+} = c\dot{\gamma}^i(a_q) \qquad (i = 1, \cdots, m). \tag{4.4}$$

Such corners will be called *tangential* corners. A broken solution of (3.9) whose subarcs are tangential solutions will have tangential corners as one sees readily. With this understood we state the following lemma.

LEMMA 4.2. *A necessary and sufficient condition that a set $(z) \neq (0)$ be a critical point of $Q(z)$ is that (z) determine a special solution of the family* (3.9).

Let (z) be a critical point of $Q(z)$. For q fixed we shall show that the n conditions

$$Q_{u_q^j} = 0 \qquad (j = 1, \cdots, n) \tag{4.5}$$

imply that $t = a_q$ is a tangential corner of the broken solution determined by (z). For the conditions (4.5) can be written more explicitly in the form (q not summed)

$$\Omega_{\dot{\eta}^i} \Big|_{a_q^-}^{a_q^+} x^i_{jq}(a_q, 0) = 0 \qquad (i = 1, \cdots, m; j = 1, \cdots, n), \tag{4.6)'}$$

as follows from (3.8). Furthermore

$$\Omega_{\dot{\eta}^i} \Big|_{a_q^-}^{a_q^+} \dot{\gamma}^i(a_q) = 0. \tag{4.6)''}$$

In fact the left member of (4.6)″ takes the form

$$(\dot{\gamma}^i F^0_{r^ir^j})\Delta\dot{\eta}^j \qquad (t = a_q; i, j = 1, \cdots, m), \tag{4.7}$$

and the parentheses in (4.7) are null. As a system of equations on the variables

$$\Omega_{\dot{\eta}^i} \Big|_{a_q^-}^{a_q^+} \qquad (q \text{ fixed}), \tag{4.8}$$

the combined equations (4.6) have a determinant (4.3) which does not vanish. Hence the variables (4.8) are null, or written more explicitly

$$F^0_{r^ir^j} \Delta\dot{\eta}^j = 0 \qquad (t = a_q; i, j = 1, \cdots, m). \tag{4.9}$$

Relations of the form (4.4) follow. The condition of the lemma is accordingly necessary.

Conversely, if (z) determines a special solution (η) in the family (3.9), then, for each q, equations (4.4), (4.9) and (4.6) are seen to hold. Hence (4.5) holds by virtue of its form (4.6)′. The set (z) defines a critical point of $Q(z)$, and the proof is complete.

If we recall that the broken solutions (η) in (3.9) are linearly independent if

and only if the sets (z) which determine them in (3.9) are linearly independent, we can infer from the preceding lemma that the nullity of $Q(z)$ equals the number of linearly independent special solutions. We come then to the proof of the following theorem.

THEOREM 4.1. *The nullity of $Q(z)$ equals the order of $t = t_2$ as a conjugate point of $t = t_1$.*

Let q be any one of the integers $1, \cdots, p$. We define a broken solution $\eta_q^i(t)$ as follows (q fixed):

$$
\begin{aligned}
\eta_q^i &\equiv (t - a_0)\dot{\gamma}^i, && a_0 \leqq t \leqq a_q, \\
\eta_q^i &\equiv k(t - a_{q+1})\dot{\gamma}^i, && a_q \leqq t \leqq a_{q+1}, \qquad (4.10)\\
\eta_q^i &\equiv 0, && a_{q+1} \leqq t \leqq a_{p+1},
\end{aligned}
$$

where k is a constant so chosen that the solution η_q^i hereby defined is continuous at $t = a_q$. The choice of k is

$$k = \frac{a_q - a_0}{a_q - a_{q+1}} < 0. \tag{4.11}$$

The solution η_q^i has a tangential corner at $t = a_q$. For at $t = a_q$ we have

$$\Delta\dot{\eta}_q^i = (k - 1)\dot{\gamma}^i,$$

and $(k - 1) \neq 0$ as follows from (4.11).

To establish the theorem we begin by showing how each special broken solution $\eta^i(t)$ which is determined in (3.9) by a critical point (z) of $Q(z)$ determines an ordinary solution of the R. J. E. which vanishes at t_1 and t_2. In fact the broken solution

$$\boldsymbol{\eta}^i = \eta^i(t) - c_q\eta_q^i(t) \tag{4.12}$$

vanishes at t_1 and t_2 and becomes an ordinary solution if the constants $c_1, \cdots, c_p$ are successively chosen so that $\boldsymbol{\eta}^i$ has no corners at the points a_q. Such a choice of the constants c_q is possible since the choice of a constant c_q so as to eliminate a corner at $t = a_q$ does not introduce a new corner for which $t < a_q$. Moreover $(\boldsymbol{\eta}) \equiv (0)$ only if $(\eta) \equiv (0)$ in (4.12). For the non-vanishing of the determinants (4.3) and the condition $(\boldsymbol{\eta}) \equiv (0)$ imply that (η) is null at each of the points $t = a_q$. Hence $(\eta) \equiv (0)$ since no point $t = a_k$ is conjugate to its successor $t = a_{k+1}$. Thus linearly independent special solutions (η) determine ordinary linearly independent solutions $(\boldsymbol{\eta})$ vanishing at t_1 and t_2.

Conversely, let $(\boldsymbol{\eta})$ be an ordinary solution of the R. J. E. which vanishes at t_1 and t_2. We seek a special solution (η) which with suitable constants c_q satisfies (4.12). That is, we seek a set (z) and constants c_q such that

$$\boldsymbol{\eta}^i(t) \equiv \frac{\partial x^i(t, 0)}{\partial z^\mu} z^\mu - c_q\eta_q^i(t) \qquad (\mu = 1, \cdots, pn). \tag{4.13}$$

We start with the point $t = a_p$ and determine c_p and the last n of the variables (z) so that (4.13) holds at $t = a_p$. This is possible since the determinant (4.3) does not vanish. We next determine c_{p-1} and the n variables of the set (z) belonging to M_{p-1} so that (4.13) holds when $t = a_{p-1}$. We continue with c_{p-2}, etc., until all the constants c_q and the variables (z) have been determined. We note that the choice of c_q and the subset of (z) belonging to M_q will not affect the equations (4.13) at the points $t = a_r > a_q$. With (z) and the constants c_q so chosen, (4.13) holds at each of the points $t = a_q$, and accordingly holds identically in t.

Moreover, linearly independent solutions $(\bar{\eta})$ determine linearly independent special broken solutions (η). To establish this it is sufficient to show that a null solution (η) satisfies (4.12) with a solution $(\bar{\eta})$ and constants c_q only if $(\bar{\eta}) \equiv (0)$. But when $(\bar{\eta})$ and (η) are both without corners, the constants c_q in (4.12) must be null. Hence $(\bar{\eta}) \equiv (0)$ as stated.

The theorem follows with the aid of Corollary 2.2.

We continue with a proof of Theorem 4.2.

THEOREM 4.2. *The index of $Q(z)$ equals the sum of the orders of the conjugate points of t_1 on $t_1 < t < t_2$.*

It will be convenient to replace g by the subarc g_b on which $t_1 \leqq t \leqq b$, where $t_1 < b \leqq t_2$. On g_b we introduce the same number of intermediate vertices as previously with

$$a_0 < \cdots < a_{p+1} = b. \tag{4.13$'$}$$

With g replaced by g_b we replace $J(z)$ and $Q(z)$ by $J^b(z)$ and $Q^b(z)$, respectively. The family of broken extremals $x^i(t, z)$ will here be denoted by $x^i(t, z, b)$. As b increases the functions $x^i(t, z, b)$ will vary continuously and retain the differentiability properties of $x^i(t, z)$ at least as long as the differences $a_s - a_{s-1}$ are less than the constant ω of (3.0).

We shall prove the theorem by proving the following statement.

(α) *The index of $Q^b(z)$ equals the sum of the orders of the conjugate points of t_1 on the interval $t_1 < t < b$.*

We continue with a proof of four lemmas.

LEMMA A. *Statement (α) is true when b lies between t_1 and the first conjugate point of t_1 on g.*

The Legendre condition (1.5) and the condition that b precede the first conjugate point of t_1 on g are sufficient for the segment $t_1 \leqq t \leqq b$ of g to afford a weak minimum to J. Hence

$$J^b(z) - J^b(0) \geqq 0$$

when (z) is sufficiently near (0). Whence $Q^b(z) \geqq 0$. But $Q^b(z)$ is non-degenerate by virtue of Theorem 4.1, so that $Q^b(z) > 0$ when $(z) \neq (0)$. This establishes Lemma A.

Corresponding to any admissible set (4.13)$'$ we now denote a_{p+1} by a_{p+2} and b, and insert a new point a_{p+1} between a_p and b. We add a new inter-

mediate manifold M_{p+1} cutting g_b at $t = a_{p+1}$. The set of parameters replacing (z) will be denoted by (ζ). They will be $n(p+1)$ in number and will be chosen so that the first np components of (ζ) form the set (z). The set (ζ) will determine a broken extremal $\mathfrak{E}(\zeta)$ as previously. The broken extremal $\mathfrak{E}(\zeta)$ will coincide with the broken extremal $x^i(t, z, b)$ for $t_1 \leqq t \leqq a_p$. For the new construction the functions replacing $J^b(z)$ and $Q^b(z)$ will be denoted by $J^b_*(\zeta)$ and $Q^b_*(\zeta)$, respectively. We continue with the following lemma.

LEMMA B. $I(Q^b) \geqq I(Q^b_*)$.

We introduce a parameter e. The inequality

$$J^b_*(e\zeta) - J^b(ez) \geqq 0 \tag{4.14}$$

holds for e sufficiently small and for the first pn components of (ζ) given by (z). Regarded as a function $\varphi(e)$, the left member of (4.14) has a minimum when $e = 0$. Hence

$$\varphi''(0) = J^b_{*\zeta^r\zeta^s}(0)\zeta^r\zeta^s - J^b_{z^\mu z^\nu}(0)z^\mu z^\nu \geqq 0, \tag{4.15}$$

$$[\mu, \nu = 1, \cdots, np; r, s = 1, \cdots, n(p+1)].$$

The lemma follows from Lemma 3.1 (a).

Statement (α) is true for b sufficiently near t_1 as we have seen in Lemma A. Let c be a constant such that $t_1 < c \leqq t_2$. Statement (α) will follow for all values of b for which $t_1 < b \leqq t_2$ once we have established the following lemmas.

LEMMA C. *If (α) is true for $b < c$, it is true for $b = c$.*

LEMMA D. *If (α) is true for $b \leqq c$, it is true for $b > c$ and sufficiently near (c).*

As b increases $I(Q^b)$ changes at most when b passes through a conjugate point a of t_1 and will then increase by at most the order ν_a of a as a conjugate point of t_1. Hence for each value of c,

$$I(Q^c) \leqq \sum \nu_a \qquad (t_1 < a < c). \tag{4.16}$$

To apply Lemma B we take $a_{p+2} = c$ and a_{p+1} nearer c than any conjugate point of t_1. We then have

$$I(Q^c) \geqq I(Q^c_*) \geqq I(Q^{a_{p+1}}). \tag{4.17}$$

The first inequality follows from Lemma B. The second follows upon putting the last n of the variables in Q^c_* equal to zero and applying Lemma 3.2 (a). But since we are assuming that statement (α) is true for $b < c$, we have

$$I(Q^{a_{p+1}}) = \sum \nu_a \qquad (t_1 < a < c).$$

It follows from (4.17) that

$$I(Q^c) \geqq \sum \nu_a \qquad (t_1 < a < c). \tag{4.18}$$

Lemma C follows from (4.16) and (4.18).

To establish Lemma D let b be a constant $> c$ but nearer c than any conjugate

point of t_1, possibly excepting c. We introduce the manifold M_{p+1} cutting g_b at $t = c$, and note that

$$(4.19) \qquad I(Q^b) \geqq I(Q^b_*) \geqq I(Q^c) + N(Q^c).$$

The first inequality follows from Lemma B. The second follows upon setting the last n of the variables in Q^b_* equal to zero and applying Lemma 3.2 (b). But $N(Q^c)$ is the order of c as a conjugate point of t_1. Hence (4.19) yields the result

$$(4.20) \qquad I(Q^b) \geqq \sum \nu_a \qquad (t_1 < a < b).$$

Lemma D follows upon comparing (4.20) with (4.16).

The proof of the theorem is complete. The Index Theorem follows from Theorems 4.1 and 4.2.

The index and nullity of $Q(z)$ depend only on the conjugate points of t_1 on g and their orders, and are accordingly independent of the number, position and representation of the manifolds M_q, provided these manifolds are admissibly distributed and represented.

INSTITUTE FOR ADVANCED STUDY.

STRUCTURES AND GROUP THEORY. II

By Oystein Ore

In this second paper on the application of the theory of structures to group theory one finds a further study of the decomposition theorems for groups. In the first chapter the completely reducible factorizations of a group are considered. Some of the theorems on normal factorizations are direct translations of the results on general Dedekind structures, but various new structural properties have also been obtained. A fundamental concept is the *normal cover,* a characteristic subgroup already introduced and studied by Remak. The duality principle gives interesting connections between the upper and lower factorizations. The *element covers* define another chain of characteristic subgroups, and the duality principle shows the correspondence between the element cover and the ϕ-group of the group.

In the second chapter the representation of a group as the union of subgroups is considered. It is shown as in the case of the Jordan-Hölder theorem that the main theorem not only holds for normal decompositions, a case considered by Kurosch, but also for mutually permutable or quasi-normal decompositions. In the normal case necessary and sufficient conditions for unique decompositions are given. A necessary condition is that the anticenter, i.e., the quotient group with respect to the commutator group, be cyclic. The dual representations by means of cross-cuts are also considered.

In the third chapter some properties of the distributive law are discussed. For finite groups all distributive pairs of subgroups are determined. It follows simply that the only finite groups in which the structure of all subgroups is distributive are the cyclic groups. Finally, a connection between distributive substructures and primitive elements is derived.

Chapter 1. Factorizations of a group into completely reducible factors

1. **Minimal groups.** In I, Chapter 1 we have introduced the concept of the product of two quotient groups.[1] In terms of this product the ordinary theorem of Jordan-Hölder takes the form of a factorization theorem about groups factorable into prime factors. In the following we shall derive other decomposition theorems where the factors are not simple but have other characteristic properties.

Let us suppose that the descending chain condition is satisfied for the normal subgroups of G. Then there exist *minimal normal subgroups* of G, i.e., normal subgroups of G which contain no other normal subgroups of G. Two minimal groups must be relatively prime when they are different.

Received March 11, 1938.

[1] O. Ore, *Structures and group theory.* I, this Journal, vol. 3(1937), pp. 149–174.

THEOREM 1. *If two different minimal groups are directly similar, they are Abelian.*

Proof. Let P and Q be the given minimal groups. In the conditions

$$[P, R] = [Q, R], \qquad (P, R) = (Q, R) = E$$

we may assume that the normal subgroup R of G is contained in $[P, Q]$. This is seen to imply that R is also a minimal group and hence we find

$$[P, Q] = [P, R] = [Q, R], \qquad (P, Q) = (P, R) = (Q, R) = E,$$

and the theorem follows from I, Chapter 4, Theorem 2.

THEOREM 2. *Let P_1, P_2, P_3 be minimal groups. If P_1 is directly similar to P_2 and P_2 to P_3, then P_1 is directly similar to P_3.*

Proof. We may assume that all the given groups are different. The conditions

$$[P_1, Q_1] = [P_2, Q_1] = [P_1, P_2], \qquad [P_2, Q_2] = [P_3, Q_2] = [P_2, P_3]$$

imply

$$[P_1, Q_1, Q_2] = [P_3, Q_1, Q_2].$$

Hence the theorem is proved when P_1 and P_3 are not contained in $[Q_1, Q_2]$. In the remaining case

$$[Q_1, Q_2] = [P_1, P_3] = [P_1, P_2] = [P_3, P_2].$$

A minimal group P of G need not be simple. Let us suppose that the chains of normal subgroups in P are finite. There exists a minimal group R of P. All the conjugates of R in G are seen to be minimal groups of P. It is easily seen that one can represent P by a basis

$$P = [R_1, R_2, \cdots, R_t], \tag{1}$$

where the R_i are some of the conjugates of R and P in (1) is the direct union of these components. Here all R_i are simple because a normal subgroup of R_i would be normal in P.

If all the conjugates of R do not occur in the direct product (1), it is easily seen that a missing R_j could replace some R_i in (1). Then R_i and R_j would be Abelian because of Theorem 1, and since they are simple, all R_i would be cyclic of the same prime order p.

THEOREM 3. *Let the normal chains in a minimal group P of G be finite. If P is non-Abelian and contains a simple normal subgroup R, it is equal to the direct product of all the conjugates of R in G. When P is Abelian, it is the direct product of conjugate cyclic groups of prime order.*[2]

[2] This theorem is due to R. Remak, *Über minimale invariante Untergruppen in der Theorie der endlichen Gruppen*, Journal für Math., vol. 162(1930), pp. 1–16. For finite groups the results of the following §2 are also due to Remak. Similar considerations can be found in K. Shoda, *Über direktzerlegbare Gruppen*, Journal of the Faculty of Science, Tokyo, Sect. 1, vol. 2(1929), pp. 51–72; *Bemerkungen über vollständig reduzible Gruppen*, ibid., vol. 2(1929), pp. 203–209.

2. **Normal covers.** The union C of all minimal subgroups of G shall be called the *normal cover* of G. It may also be defined as the smallest normal subgroup with which every normal subgroup of G has a common subgroup.

This second definition gives

THEOREM 4. *The normal cover is a characteristic subgroup of G.*

Let us now suppose that all normal chains in G are finite. Then there exists a basis representation of C as the direct product of certain of the minimal groups of G

$$C = [P_1, P_2, \cdots, P_r]. \tag{2}$$

It is now obvious from Theorem 3 that we have

THEOREM 5. *If the normal chains in the minimal groups of G are finite, the normal cover C of G is the direct product of simple groups.*

From the general theory of Dedekind structures[3] follows

THEOREM 6. *Any two basis representations*

$$C = [P_1, \cdots, P_2] = [Q_1, \cdots, Q_2] \tag{3}$$

of the normal cover contain the same number of minimal groups directly similar in pairs.

One can also prove this theorem directly by showing that each basis element P_i in one basis representation (3) can replace some suitably chosen Q_j in the other. One also sees that any normal subgroup D of G which is contained in C has a basis representation

$$D = [P_1, P_2, \cdots, P_s],$$

which can be completed into a basis for C

$$C = [D, P_{s+1}, \cdots, P_r].$$

All of these facts follow since the structure of normal subgroups of G contained in C is a *completely reducible structure.*

Two minimal groups which are directly similar shall be said to belong to the same *uniform system.* Since the concept of direct similarity is transitive for minimal groups according to Theorem 2, all minimal groups fall into distinct uniform systems. The union of all minimal groups in such a system shall be called a *maximal uniform component.* It is then easy to prove

THEOREM 7. *The normal cover is representable uniquely as the direct union of its maximal uniform groups*

$$C = [U_1, \cdots, U_k].$$

This representation may be obtained from any basis representation of C by joining the basis elements belonging to the same uniform system into one term. A non-Abelian minimal group is its own maximal uniform group.

[3] See O. Ore, *On the foundation of abstract algebra.* II, Annals of Math., vol. 37(1936), pp. 265–292. The results which are used in the following can be found in Chapter 3.

3. **Product representations.** To insure uniform notation in the following we shall write every group as a quotient group with respect to the unit element. We write in particular

$$\mathfrak{G} = G/E, \qquad \mathfrak{M}_1 = C/E, \qquad \mathfrak{G} = \mathfrak{G}_1 \times \mathfrak{M}_1 .$$

Now the group $\mathfrak{G}_2 = G/C$ has itself a normal cover $\mathfrak{C}_2 = C_2/C_1$. We shall call C_2 the *second normal cover* of G, and $\mathfrak{C}_2$ the *second normal cover quotient.* Continuing, we find

THEOREM 8. *A group $\mathfrak{G}$ in which the finite chain condition holds for normal subgroups has a unique factorization*

$$\mathfrak{G} = \mathfrak{C}_h \times \cdots \times \mathfrak{C}_2 \times \mathfrak{C}_1 \tag{4}$$

as the product of its successive normal cover quotients.

We also easily see the truth of

THEOREM 9. *The successive covers $C = C_1, C_2, \cdots$ are characteristic subgroups of G.*

Let M and N be two normal subgroups of G such that $M \geqq N$. We shall say that the quotient $\mathfrak{M} = M/N$ is *completely reducible* in G if

$$\mathfrak{M} = [\mathfrak{Q}_1, \cdots, \mathfrak{Q}_s]$$

is the direct union of minimal groups $\mathfrak{Q}_i = Q_i/N$ in G/N. From Theorem 3 it follows that $\mathfrak{M}$ is the direct product of simple groups, when descending normal chains are finite in all $\mathfrak{Q}_i$.

The general theory of structures now gives the following property of the factorization (4):

THEOREM 10. *Let $\mathfrak{H}$ be a normal subgroup of $\mathfrak{G}$ and*

$$\mathfrak{H} = \mathfrak{H}_t \times \cdots \times \mathfrak{H}_2 \times \mathfrak{H}_1$$

some factorization of $\mathfrak{H}$ into completely reducible factors. If $\mathfrak{G}$ has the factorization (4) *into its cover quotients, for each i the quotient $\mathfrak{C}_i \times \cdots \times \mathfrak{C}_1$ has the factor $\mathfrak{H}_i \times \cdots \times \mathfrak{H}_1$.*

A consequence of this theorem is

THEOREM 11. *Any representation of a group $\mathfrak{G}$ as the product of completely reducible factors contains at least as many factors as the cover factorization* (4).

This theorem shows that among the completely reducible factorizations of $\mathfrak{G}$ the cover factorization is one of shortest length. This property, however, is not sufficient to characterize the cover factorization completely.

Let $H = \mathfrak{H}$ be a normal subgroup of $\mathfrak{G}$ and let us call the union of all minimal groups of $\mathfrak{G}$ contained in $\mathfrak{H}$ the first normal cover $\mathfrak{H}_1$ of $\mathfrak{H}$ with respect to $\mathfrak{G}$. The second normal cover is defined as the union of those normal subgroups of $\mathfrak{G}$ which are minimal over $\mathfrak{H}_1$ and contained in $\mathfrak{H}$, etc. By repetition of this process one obtains a unique factorization

$$\mathfrak{H} = \mathfrak{H}_{h_1} \times \cdots \times \mathfrak{H}_1 \tag{5}$$

of $\mathfrak{H}$ as the product of its successive cover quotients with respect to $\mathfrak{G}$.

THEOREM 12. *Let $\mathfrak{H}$ be a normal subgroup of $\mathfrak{G}$. The cover factorization of $\mathfrak{H}$ with respect to $\mathfrak{G}$ is the same as the cover factorization of $\mathfrak{H}$ in itself.*

Proof. Let us show first that $\mathfrak{H}_1$ in (5) is the cover of $\mathfrak{H}$. Let P be a minimal group of $\mathfrak{G}$ contained in $\mathfrak{H}$ and R some minimal group of $\mathfrak{H}$ contained in P. One sees as before that

$$P = [R_1, \cdots, R_t]$$

is the direct product of some of the conjugates of R in $\mathfrak{G}$. These conjugates

$$R_i = g_i R g_i^{-1}$$

are all normal and minimal in $\mathfrak{H}$ since $\mathfrak{H}$ is normal in $\mathfrak{G}$. Since P is the union of minimal groups of $\mathfrak{H}$, it is contained in the cover of $\mathfrak{H}$. Conversely, a minimal group R of $\mathfrak{H}$ must be contained in a minimal group P of $\mathfrak{G}$ because any normal subgroup of $\mathfrak{G}$ containing one of the conjugates of R must contain them all.

This shows that $\mathfrak{H}_1$ in (5) is the cover of $\mathfrak{H}$. When the same argument is applied to the successive factor groups $\mathfrak{H}/\mathfrak{H}_i$ and $\mathfrak{G}/\mathfrak{H}_i$, the theorem follows.

THEOREM 13. *Let $\mathfrak{H}$ be a normal subgroup of $\mathfrak{G}$ and* (4) *and* (5) *their cover factorization. Then for each i*

$$\mathfrak{H}_i \times \cdots \times \mathfrak{H}_1 = (\mathfrak{H}, \mathfrak{C}_i \times \cdots \times \mathfrak{C}_1).$$

Proof. The preceding proof shows that the theorem is true for $i = 1$. We shall prove the theorem by induction and hence we can assume

$$\mathfrak{H}_{i-1} \times \cdots \times \mathfrak{H}_1 = (\mathfrak{H}, \mathfrak{C}_{i-1} \times \cdots \times \mathfrak{C}_1).$$

Let us write

$$\mathfrak{C}_{i-1} \times \cdots \times \mathfrak{C}_1 = \mathfrak{R} \times \mathfrak{H}_{i-1} \times \cdots \times \mathfrak{H}_1 .$$

We then find

$$\mathfrak{C}_h \times \cdots \times \mathfrak{C}_i \times \mathfrak{R} = \mathfrak{C} \times \mathfrak{H}_{h_1} \times \cdots \times \mathfrak{H}_i ,$$

and here $\mathfrak{R}$ must be relatively prime to $\mathfrak{H}_i$ and $\mathfrak{H}_{h_1} \times \cdots \times \mathfrak{H}_i$ because otherwise $\mathfrak{H}_{i-1} \times \cdots \times \mathfrak{H}_1$ would not be the greatest common factor of $\mathfrak{H}$ and $\mathfrak{C}_{i-1} \times \cdots \times \mathfrak{C}_1$. From the last relation it is easily seen that

$$(\mathfrak{C}_i , \mathfrak{R}^{-1}(\mathfrak{H}_{h_1} \times \cdots \times \mathfrak{H}_i)\mathfrak{R}^{-1}) = \mathfrak{R}\mathfrak{H}_i\mathfrak{R}^{-1}.$$

This relation gives in turn after multiplication with $\mathfrak{R} \times \mathfrak{H}_{i-1} \times \cdots \times \mathfrak{H}_1$

$$(\mathfrak{C}_i \times \cdots \times \mathfrak{C}_1 , [\mathfrak{H}, \mathfrak{C}_{i-1} \times \cdots \times \mathfrak{C}_1]) = [\mathfrak{H}_i \times \cdots \times \mathfrak{H}_1 , \mathfrak{C}_{i-1} \times \cdots \times \mathfrak{C}_1].$$

By taking the cross-cut of this relation by $\mathfrak{H}$ one obtains the desired result (6).

It may be observed that Theorem 13 is a general structure theorem valid for all Dedekind structures.

The number h of successive covers of $\mathfrak{G}$ in (4) may be called the *normal height* of $\mathfrak{G}$. One can also define the normal height h_H with respect to $\mathfrak{G}$ of any subgroup H as the smallest index i such that $\mathfrak{C}_i \times \cdots \times \mathfrak{C}_1$ contains H.

This height has the evaluation properties that the height of the union is equal to the maximal height of any of the components while the height of a cross-cut is at most equal to the minimal height among the groups of which one takes the cross-cut. Theorems 12 and 13 show that for a normal subgroup $\mathfrak{H}$ the height is independent of the group $\mathfrak{G}$ in which $\mathfrak{H}$ is normal, and it is equal to the height of $\mathfrak{H}$ with respect to itself.

4. **Upper covers.** Several results of interest may be obtained by considering the duals of the preceding theorems. We suppose as before that the ascending and descending chain condition is satisfied for normal subgroups in G. A *maximal normal group* M in G is one having the property that there are no normal subgroups between G and M. The corresponding quotient group $\mathfrak{P} = G/M$ shall be called an *upper* (or *left-hand*) *minimal factor* of $\mathfrak{G} = G/E$. We obviously have

THEOREM 14. *An upper minimal factor is simple, and hence non-Abelian or cyclic of prime order.*

The left-hand union $\mathfrak{U}$ of all minimal upper factors shall be called the *upper cover quotient* of $\mathfrak{G}$. If D is the cross-cut of all maximal groups of $\mathfrak{G}$, then $\mathfrak{D} = G/D$, and we shall call D the *upper cover group* of $\mathfrak{G}$. We can write $\mathfrak{D}$ as the left-hand direct union of upper minimal factors

$$\mathfrak{D} = [\mathfrak{P}_1, \cdots, \mathfrak{P}_k]_l,$$

or according to Theorem 6, Chapter 1, I as a direct right-hand union or product

$$\mathfrak{D} = [\mathfrak{Q}_1, \cdots, \mathfrak{Q}_k], \qquad \mathfrak{Q}_i = Q/D, \tag{7}$$

where the quotient group $\mathfrak{Q}_i$ is similar to $\mathfrak{P}_i$.

We can define the second upper normal cover group D_2 as the upper cover group of $D = D_1$ and the second upper cover quotient as $\mathfrak{D}_2 = D_1/D_2$, etc. By the definition of the second and following cover groups we have the choice of taking the maximal normal subgroups of D_1 or taking those normal subgroups of G which are maximal in D_1. It is, however, easily seen that both definitions lead to the same result. We find as before

THEOREM 15. *There exists a unique factorization of $\mathfrak{G}$ by means of its successive upper cover quotients*

$$\mathfrak{G} = \mathfrak{D}_1 \times \cdots \times \mathfrak{D}_h.$$

We also easily obtain

THEOREM 16. *The upper cover groups of $\mathfrak{G}$ are characteristic subgroups.*

We shall finally consider the relation between the upper and the preceding lower cover factorization. It follows from (7) that the upper cover factorization is completely reducible. According to Theorem 11 this implies that the upper factorization does not contain fewer factors than the lower. Since the dual argument is also valid we have

THEOREM 17. *The upper and lower cover factorizations*

$$(8)\qquad \mathfrak{G} = \mathfrak{D}_h \times \cdots \times \mathfrak{D}_1 = \mathfrak{C}_h \times \cdots \times \mathfrak{C}_1$$

contain the same number of factors. For each i the quotient $\mathfrak{C}_i \times \cdots \times \mathfrak{C}_1$ has the right-hand factor $\mathfrak{D}_i \times \cdots \times \mathfrak{D}_1$ and $\mathfrak{D}_h \times \cdots \times \mathfrak{D}_i$ has the left-hand factor $\mathfrak{C}_h \times \cdots \times \mathfrak{C}_i$.

The last statement is a consequence of Theorem 10 and its dual.

Let us call any other completely reducible factorization of $\mathfrak{G}$ with the same number h of factors as the cover factorizations (8) a shortest *completely reducible factorization.* It is then easily seen that one can characterize the cover factorizations in the set of all shortest factorizations in the following manner.

THEOREM 18. *Let the upper and lower cover factorizations of $\mathfrak{G}$ be given by* (8) *and let*

$$\mathfrak{G} = \mathfrak{H}_h \times \cdots \times \mathfrak{H}_1$$

be an arbitrary shortest completely reducible factorization. Then for each i the quotient $\mathfrak{H}_i \times \cdots \times \mathfrak{H}_1$ is a right-hand factor of $\mathfrak{C}_i \times \cdots \times \mathfrak{C}_1$ and has the right-hand factor $\mathfrak{D}_i \times \cdots \times \mathfrak{D}_1$; and dually $\mathfrak{H}_h \times \cdots \times \mathfrak{H}_i$ is a left-hand factor of $\mathfrak{D}_h \times \cdots \times \mathfrak{D}_i$ and has the left-hand factor $\mathfrak{C}_h \times \cdots \times \mathfrak{C}_i$.

5. Element covers. To conclude, let us mention a different type of covers. Let us consider all subgroups of G instead of only the normal subgroups as in the preceding. There exist in G subgroups F such that every subgroup of G different from the unit element E has a subgroup (or an element) E in common with F. We can prove: *the set of all such groups forms a structure.* It is obvious that the union of two such groups F_1 and F_2 has the same property. To prove it for the cross-cut (F_1, F_2), let D be a subgroup having a subgroup D_1 in common with F_1. Then D_1 has a common subgroup D_2 with F_2. Hence D_2 is a common subgroup of D and (F_1, F_2).

If there exists a minimal group F_1 of G with the property that every subgroup has a common subgroup with F_1, then F_1 shall be called the *element cover* of G. One easily sees the truth of

THEOREM 19. *The element cover is a characteristic subgroup of G.*

When the element cover exists, it contains the normal cover C_1 because the latter was defined as the normal group having a subgroup in common with every normal subgroup. The element cover of G must contain all (non-normal) minimal groups, i.e., all elements of prime order. Hence we have

THEOREM 20. *Let G be a group in which every element is of finite order. Then the element cover exists and is the group generated by all elements of prime order.*

As before, we may define the second element cover F_2 as the first element cover of G/F_1 and the second element cover quotient as $\mathfrak{C}_2 = F_2/F_1$.

The second element cover F_2 is generated by all those elements f_2 of G whose order over F_1 is a prime p_i, i.e., $f_2^{p_i}$ can be expressed by means of products of elements of prime order. When this process is continued, one finds

THEOREM 21. *Every group is representable uniquely as the product of its successive element cover quotients. The successive element cover groups are characteristic subgroups.*

As usual one can construct the dual concepts. The dual or upper element cover must be defined as the cross-cut of all maximal subgroups of G. This is, however, exactly the definition of the well-known ϕ-group of G. One can also define the ϕ-group structurally in a different manner. We consider all representations of G as the union of subgroups

$$G = [A_1, \cdots, A_r].$$

All those subgroups A of G which are superfluous in any such representation are seen to form a structure. The union $\bar{F}$ of all such groups is the ϕ-group of G. Let us suppose that there existed a maximal subgroup M of G not containing $\bar{F}$. Then one would have

$$G = [\bar{F}, M],$$

and $\bar{F}$ would not be superfluous. This shows that $\bar{F}$ is contained in every maximal subgroup of G, so that $\phi \supseteqq \bar{F}$. Now let us consider an element d not contained in $\bar{F}$. There exists a representation

$$G = [\{d\}, K],$$

where the cyclic group $\{d\}$ cannot be omitted. Among the various K let us consider a maximal one. Then K is also a maximal group of G because the addition of any element to K would make $\{d\}$ superfluous. Hence we have $\phi = \bar{F}$.

This property can be dualized to the element cover. We consider all representations of G as a left-hand union, i.e., all representations of the unit element E as the cross-cut of subgroups

$$(A_1, \cdots, A_r) = E. \tag{9}$$

Again, one sees that those subgroups which can be omitted in all such representations form a structure and that there exists a minimal group $\bar{\phi}$ such that $\bar{\phi}$ and any group containing $\bar{\phi}$ can always be omitted. We can then show that this group is identical with the element cover group F_1. Obviously, any group which can always be omitted must have the property that

$$(\phi, \{c\}) \neq E$$

for any cyclic subgroup $\{c\}$ of G. This implies that $\bar{\phi} \supseteqq F_1$. On the other hand, if

$$(F_1, A_1, \cdots, A_r) = E,$$

one must also have the relation (9).

This dualism raises a very interesting problem. The structure of the ϕ-group is known to possess the simple property that it is the direct product of its

Sylow groups.[4] The question then arises whether it is possible to make any general statement about the structural properties of the dual group, namely, the quotient group G/F.

Chapter 2. Decomposition theorems

1. **Components.** Any representation of a group G as the union of subgroups

$$G = [A_1, A_2, \cdots, A_n] \tag{1}$$

shall be called a *decomposition* of G. The existence of a decomposition (1) indicates that G can be generated by the elements of the subgroups A_i. We shall write

$$\bar{A}_i = [A_1, \cdots, A_{i-1}, A_{i+1}, \cdots, A_n], \tag{2}$$

and call $\bar{A}_i$ the *complement* of A_i in the decomposition (1).

We shall call (1) a *permutable* decomposition when each group A_i is permutable with all other A_j. It follows that A_i is also permutable with its complement $\bar{A}_i$. In this case, every element g of G can be represented in the form

$$g = a_1 \cdot a_2 \cdots a_n, \tag{3}$$

where a_i belongs to A_i.

Now let D be a subgroup of G. The group

$$D_i = (A_i, [D, \bar{A}_i]) \tag{4}$$

we shall call the *component of D in A_i*. For permutable decompositions these components have various properties which we shall now indicate.

Lemma 1. *The component D_i is permutable with all A_j.*

This is a consequence of Lemma 3, Chapter 2, I.

Lemma 2. *The component D_i is permutable with all other components D_j.*

Since $[D, \bar{A}_j] > D_i$, this result follows by the same Lemma 3, Chapter 2, I.

When the elements d of D are written in the form (3), all elements a_i occurring in the product shall be called the *representatives* of D in A_i in the decomposition (1). We can then prove:

Lemma 3. *The component D_i contains all representatives of D in A_i.*

Proof. We write

$$d = a_i \cdot \bar{a}_i .$$

Hence any representative a_i belongs both to A_i and to $[D, \bar{A}_i]$. It is easily seen that D_i cannot contain any proper subgroup which is permutable with $\bar{A}_i$ and contains all representatives of D. Usually not every element of D_i is a representative. We can prove however:

Lemma 4. *When D is permutable with $\bar{A}_i$, then D_i is the set of all representatives of D in A_i.*

[4] Miller, Blichfeldt and Dickson, *Theory and Application of Finite Groups*, p. 72.

Proof. Since D is permutable with $\bar{A}_i$, all elements of $[D, \bar{A}_i]$ are of the form $d \cdot \bar{a}_i$, and all elements a_i of D_i satisfy a relation

$$a_i = d \cdot \bar{a}_i, \qquad d = a_i \cdot \bar{a}_i^{-1}.$$

Hence a_i is a representative.

Lemma 3 obviously implies

LEMMA 5. *The group D is contained in the union of its components in the various A_i:*

$$D < [D_1, D_2, \cdots, D_n].$$

We shall also need the following

LEMMA 6. *Let C and D be two subgroups of G and C_i and D_i their components in A_i. Then the union $[C, D]$ has the component $[C_i, D_i]$ in A_i.*

Proof. According to Lemma 1 the union $[C_i, D_i]$ is permutable with A_i, and hence we find from Theorem 6, Chapter 2, I,

$$[B_i, C_i] = (A_i, [B_i, C_i, \bar{A}_i]). \tag{5}$$

Now we find from the same theorem

$$[\bar{A}_i, B_i] = [\bar{A}_i, (A_i, [\bar{A}_i, B])] = ([\bar{A}_i, B], [A_i, \bar{A}_i]) = [\bar{A}_i, B],$$
$$[\bar{A}_i, C_i] = [\bar{A}_i, C],$$

and when this is substituted in (5), we obtain

$$[B, C]_i = (A_i, [B, C, \bar{A}_i]) = [B_i, C_i].$$

Lemma 6 can of course be extended to an arbitrary number of groups and this leads to the following important result.

LEMMA 7. *Let* (1) *be a permutable decomposition and*

$$G = [B_1, \cdots, B_m] \tag{6}$$

an arbitrary decomposition of G. When $A_{i,j}$ is the component of B_j in A_i we have

$$A_i = [A_{i,1}, \cdots, A_{i,m}]. \tag{7}$$

We shall prove finally

LEMMA 8. *Let B and C be permutable and let both be permutable with $\bar{A}_i$. Then the components B_i and C_i are also permutable.*

Proof. The elements of B and C can be written in the form

$$b = a_i^{(b)} \cdot \bar{a}_i^{(b)}, \qquad c = a_i^{(c)} \cdot \bar{a}_i^{(c)},$$

where $a_i^{(b)}$ and $a_i^{(c)}$ run through B_i and C_i, respectively. Since B and C are permutable, we have $b \cdot c = c_1 \cdot b_1$, and one finds that the representatives of $[B, C]$ may be written both in the form $a_i^{(b)} \cdot a_i^{(c_1)}$ and in the form $a_i^{(c)} \cdot a_i^{(b_1)}$.

2. Permutable and quasi-normal decompositions. We shall say that a decomposition (1) is a *permutable decomposition into indecomposable parts* when

no group A_i can be further decomposed so that the resulting decomposition is permutable. In such a decomposition we suppose further that no term A_i is *superfluous*, i.e., is contained in its complement $\bar{A}_i$. Finally, we say that two permutable decompositions (1) and (6) are *mutually permutable* when each A_i is permutable with each B_j. We can then prove as the main theorem

THEOREM 1. *Let*

$$G = [A_1, \cdots, A_n] = [B_1, \cdots, B_m] \tag{8}$$

be two mutually permutable decompositions of a group G into indecomposable parts without superfluous terms. Then both decompositions have the same number of terms and any term in one decomposition may be replaced by a suitably chosen term in the other to give a new decomposition of the same kind.

Proof. Under the given assumptions it follows that in (7) one of the A_i must be equal to $A_{i,j}$. Let us suppose that

$$A_1 = A_{1,1} = (A_1, [B_1, \bar{A}_1]).$$

This implies

$$G = [B_1, A_2, \cdots, A_n], \tag{9}$$

and shows that A_1 can be replaced by B_1. Possibly some term in (9) might be superfluous. Let us suppose for the moment that $n \leqq m$. We continue the process and replace all A_i by B_j. Since no B_j is superfluous in (8), we find that the first decomposition (8) must contain at least as many as the second and that no term in (9) can be superfluous. Hence $n = m$, and the theorem is proved.

In a similar manner, we can prove a decomposition theorem for a group into quasi-normal components.

THEOREM 2. *Let there exist two representations without superfluous terms*

$$G = [A_1, \cdots, A_n] = [B_1, \cdots, B_m]$$

of a group G as the union of quasi-normal subgroups. Each A_i and B_j shall be quasi-normally indecomposable, i.e., A_i is not representable as the union of two proper quasi-normal subgroups of A_i. Then both decompositions contain the same number of terms and any term in one decomposition may be replaced by a suitably chosen term in the other to give a new decomposition of the same kind.

Proof. Since the union of two quasi-normal subgroups is again quasi-normal, we have that $[\bar{A}_i, B_j]$ is quasi-normal in G. From Theorem 15, Chapter 2, I it follows that $A_{i,j}$ in (7) is quasi-normal in A_i. Hence we conclude as before that one $A_{i,j}$ must be equal to A_i and we can assume that

$$G = [B_1, \bar{A}_1].$$

From this point on the proof is analogous to the preceding.

3. Normal decompositions. We shall next turn to the representation of groups as the union of normal subgroups. We shall say that a normal sub-

group A of G is *indecomposable* with respect to G when it is not the union of two proper subgroups both normal in G. We then find as before[5]

THEOREM 3. *Let*

$$G = [A_1, \cdots, A_n] = [B_1, \cdots, B_m] \tag{10}$$

be two decompositions without superfluous terms, where the normal subgroups A_i and B_j are indecomposable with respect to G. Both decompositions then contain the same number of terms and any term in one decomposition can be replaced by some suitably chosen term of the other.

The proof is based on Lemma 7 as before. In this case, all $A_{i,j}$ are normal subgroups of G. Let us mention without proof that one has the stronger replacement theorem:

THEOREM 4. *Any term A_i in* (10) *can be interchanged with some B_j such that*

$$G = [A_i, \bar{B}_j] = [B_j, \bar{A}_i].$$

If we say that the first decomposition (10) is reduced when no group A_i can be replaced by a proper subgroup also normal in G, we can state

THEOREM 5. *The decompositions of Theorem* 3 *are reduced.*

Proof. If one had for $A_i' < A_i$

$$[A_i, \bar{A}_i] = [A_i', \bar{A}_i],$$

one would obtain from the Dedekind law

$$A_i = [A_i', (A_i, \bar{A}_i)],$$

and A_i would not be indecomposable with respect to G.

Let us now study somewhat further the normal subgroups of G which are indecomposable with respect to G. When we assume that the ascending chain condition holds for the normal subgroups of G, we find the following criterion:

THEOREM 6. *The necessary and sufficient condition that a normal subgroup A of G be indecomposable with respect to G is that A contain a single maximal subgroup normal in G.*

Let this maximal normal subgroup be M. Since A is minimal over M in the group G/M, it follows from Chapter 1 that the quotient group $\mathfrak{P} = A/M$ is the direct product of simple groups. We shall say that the indecomposable group A *belongs* to the quotient group $\mathfrak{P}$. We prove further

THEOREM 7. *Any indecomposable group A_i which occurs in a normal decomposition* (10) *of G belongs to a simple quotient group $\mathfrak{P}_i$.*

Proof. One establishes immediately the isomorphism

$$\mathfrak{P}_i = A_i/M_i \cong G/[M_i, \bar{A}_i],$$

and one finds that $[M_i, \bar{A}_i]$ is maximal in G. Hence the quotient groups are simple.

[5] See A. Kurosch, *Durchschnittsdarstellungen mit irreduziblen Komponenten in Ringen und in sogenannten Dualgruppen*, Matematičeski Sbornik, vol. 42(1935), pp. 613–616.

This theorem leads to the division of the indecomposable groups in the decompositions of G into two classes. This distinction is very useful in the sequel. We say that an indecomposable group A is of *Abelian type*, when its quotient group $\mathfrak{P} = A/M$ is cyclic of prime order. Otherwise, A is said to be of *non-Abelian type*.

4. **Relations between decompositions and upper covers.**[6] From now on let us suppose that all chains of normal subgroups in G have a finite length. Then there always exists a decomposition

$$G = [A_1, \cdots, A_n] \tag{11}$$

into normal indecomposable subgroups A_i. Since each A_i is indecomposable, it contains a unique normal subgroup M_i of G maximal in A_i, and according to our terminology A_i belongs to the simple group

$$\mathfrak{P}_i = A_i/M_i. \tag{12}$$

The groups

$$C_i = [M_i, \bar{A}_i] \tag{13}$$

are maximal normal subgroups of G and the quotient group

$$\mathfrak{Q}_i = G/C_i \tag{14}$$

is isomorphic to $\mathfrak{P}_i$. These quotients $\mathfrak{Q}_i$ are minimal left-hand factors of $\mathfrak{G} = G/E$. Their left-hand union is

$$\mathfrak{C} = [\mathfrak{Q}_1, \cdots, \mathfrak{Q}_n]_l = G/C, \tag{15}$$

where C is the cross-cut of all C_i. A repeated application of the Dedekind relation shows that

$$C = (C_1, \cdots, C_n) = [M_1, \cdots, M_n].$$

The quotient (15) is completely reducible. It can also be written as a right-hand union of simple groups

$$\mathfrak{C} = [\mathfrak{R}_1, \cdots, \mathfrak{R}_n],$$

where

$$R_i = N_i/C, \qquad N_i = [M_i, \cdots, M_{i-1}, A_i, M_{i+1}, \cdots, M_n],$$

and $\mathfrak{R}_i$ is similar to $\mathfrak{P}_i$. It is also easily seen that the $\mathfrak{Q}_i$ or $\mathfrak{R}_i$ form a basis for $\mathfrak{C}$.

We shall now show that $\mathfrak{C}$ is the upper cover quotient of G. If this were not true, there would exist a left-hand factor $\mathfrak{P} = G/P$ of $\mathfrak{G}$ which was rela-

[6] Compare the structure theorems in O. Ore, *On the foundation of abstract algebra*. II, Chapter 5.

tively prime to $\mathfrak{C}$. We may assume that P is maximal in G. Since it does not contain C, there is at least one M_i not contained in P. Since P is maximal

$$G = [P, M_i],$$

and this implies by the Dedekind relation

$$A_i = [M_i, (P, A_i)].$$

This is, however, contrary to the assumption that A_i is indecomposable. Hence we have

THEOREM 8. *Let G be a group in which all chains of normal subgroups are finite. There exist normal decompositions of G into groups indecomposable in G*

$$G = [A_1, \cdots, A_n]. \tag{16}$$

Any such decomposition without superfluous terms contains the same number of groups as a basis representation

$$\mathfrak{C} = [\mathfrak{R}_1, \cdots, \mathfrak{R}_n] \tag{17}$$

of the upper cover factor of G and each A_i belongs to a simple quotient group $\mathfrak{P}_i$ similar to $\mathfrak{R}_i$.

This theorem shows that the properties of the normal decompositions are mainly determined by the upper cover.

It is also easily verified by means of an extension of the preceding argument that any indecomposable A_i occurring in a decomposition (16) not only is indecomposable in G but must also be indecomposable in itself, i.e., it is not equal to the union of two proper normal subgroups of itself.

5. **Unique decompositions.** We shall now consider the conditions for the uniqueness of the normal decompositions. We prove first

THEOREM 9. *In a normal decomposition of a group into indecomposable groups those terms which are of non-Abelian type are unique.*

The proof of this theorem depends on the following

LEMMA 9. *Let A, B and C be normal subgroups of G such that*

$$[A, C] = [B, C].$$

We write

$$D_C = (A, B), \qquad D_B = (A, C), \qquad D_A = (B, C).$$

Then we have the isomorphism

$$\mathfrak{A} = A/[D_C, D_B] \cong B/[D_C, D_A] = \mathfrak{B},$$

and both quotient groups are Abelian.

Proof. One finds easily

$$\mathfrak{A}_1 = [A, D_A]/[D_A, D_B, D_C] \cong \mathfrak{A},$$

$$\mathfrak{B}_1 = [B, D_B]/[D_A, D_B, D_C] \cong \mathfrak{B}.$$

Furthermore, let

$$\mathfrak{C}_1 = [C, D_C]/[D_A, D_B, D_C].$$

Then one finds

$$[\mathfrak{A}_1, \mathfrak{C}_1] = [\mathfrak{B}_1, \mathfrak{C}_1], \qquad (\mathfrak{A}_1, \mathfrak{B}_1) = (\mathfrak{A}_1, \mathfrak{C}_1) = (\mathfrak{B}_1, \mathfrak{C}_1) = \mathfrak{E}_0,$$

where $\mathfrak{E}_0$ is a unit group. This last relation proves our lemma according to Theorem 2, Chapter 4, I.

With the same notation one also has

LEMMA 10. *If*

$$[A, C] = [B, C] = [A, B],$$

then

$$\mathfrak{A}_1 \cong \mathfrak{B}_1 \cong \mathfrak{C}_1$$

are Abelian groups and

$$[\mathfrak{A}_1, \mathfrak{C}_1] = [\mathfrak{B}_1, \mathfrak{C}_1] = [\mathfrak{A}_1, \mathfrak{B}_1], \qquad (\mathfrak{A}_1, \mathfrak{C}_1) = (\mathfrak{B}_1, \mathfrak{C}_1) = (\mathfrak{A}_1, \mathfrak{B}_1) = \mathfrak{E}_0.$$

To prove Theorem 9, let us recall that we have already shown that in a decomposition (10) any group A_i can be replaced by some B_j. Let us suppose that A_i is of non-Abelian type and

$$[A_1, \bar{A}_1] = [B_1, \bar{A}_1].$$

Since A_1 cannot have any Abelian quotient group, we conclude from Lemma 9 that

$$A_1 = [(A_1, B_1), (A_1, \bar{A}_1)],$$

and since A_1 is indecomposable, this implies $B_1 = A_1$.

Obviously, groups A_i of Abelian type can only occur when the *commutator group* K of G is a proper subgroup, i.e., when G has an *anticenter* G/K different from the unit group. Hence we can state

THEOREM 10. *In a group without anticenter the decomposition into normal indecomposable groups is unique when it exists.*

When G has an anticenter, the question of uniqueness is more difficult. Let us assume that the finite chain condition is satisfied. We can then prove

THEOREM 11. *When the decomposition of G is unique, the anticenter of G is cyclic.*

Proof. When the anticenter is not cyclic, the upper cover quotient of G must contain two different cyclic groups of the same prime order p. Hence there exists in (16) two indecomposable groups A_1 and A_2 belonging to cyclic groups

$$\mathfrak{P}_1 = A_1/M_1, \qquad \mathfrak{P}_2 = A_2/M_2.$$

If we write $C = [M_1, M_2]$, then

(18) $\quad [A_1, A_2]/C = [\mathfrak{S}_1, \mathfrak{S}_2], \qquad \mathfrak{S}_1 = [A_1, M_2]/C, \qquad \mathfrak{S}_2 = [A_2, M_1]/C,$

and $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are similar to $\mathfrak{P}_1$ and $\mathfrak{P}_2$. Since the quotient group (18) is Abelian of type (p, p), there exists a third cyclic group $\mathfrak{S} = M/C$ such that M is normal in G and

$$[\mathfrak{S}_1, \mathfrak{S}_2] = [\mathfrak{S}_1, \mathfrak{S}] = [\mathfrak{S}_2, \mathfrak{S}], \qquad (\mathfrak{S}_1, \mathfrak{S}_2) = (\mathfrak{S}_2, \mathfrak{S}) = (\mathfrak{S}_2, \mathfrak{S}),$$

or

$$[A_1, A_2] = [M, A_1] = [M, A_2].$$

Here M does not contain A_1 or A_2. By replacing M by a proper subgroup one obtains, therefore, two different indecomposable representations of $[A_1, A_2]$.

In the remaining case, where the anticenter of G is cyclic, the decomposition may be unique or not. In order to state a general theorem let us introduce the following definition: two normal subgroups A and B are said to be *directly semi-similar* in G when there exists a third normal subgroup C not containing A or B such that

$$[A, C] = [B, C].$$

We shall prove

THEOREM 12. *The necessary and sufficient condition that a decomposition* (16) *be unique is that no A_i be directly semi-similar to any normal subgroup A_j' of G contained in any other A_j.*

Proof. To prove the necessity, let us suppose that A_i and A_j' are directly semi-similar

$$[A_i, C] = [A_j', C].$$

Here C does not contain A_i, but we can suppose that it is contained in $[A_i, A_j']$. Hence

$$[A_i, A_j] = [C, A_j].$$

After C is replaced by a proper subgroup this gives a new decomposition for G.

To prove the sufficiency, we suppose that some A_i, for example A_1, can be replaced by some indecomposable B satisfying

$$[A_1, \bar{A}_1] = [B, \bar{A}_1].$$

As before, the group A_1 belongs to the group $A_1/M_1 = \mathfrak{P}_1$ of prime order p. Let us form the quotient group

$$G/M_1 = [\mathfrak{P}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_n] = [\mathfrak{B}, \mathfrak{A}_2, \cdots, \mathfrak{A}_n],$$

where

$$\mathfrak{B} = [B_1, M_1]/M_1, \qquad \mathfrak{A}_i = [A_i, M_1]/M_1.$$

Since $\mathfrak{B}$ cannot contain $\mathfrak{P}_1$, one of the $\mathfrak{A}_i$, for example $\mathfrak{A}_2$, must contain a normal cyclic factor $\mathfrak{P}_2' \cong \mathfrak{P}_1$. It is easily seen that $\mathfrak{P}_2'$ may be written in the form

$$\mathfrak{P}_2' = [A_2', M_1]/M_1,$$

where A_2' is a subgroup of A_2. Since

$$[\mathfrak{P}_1, \mathfrak{M}] = [\mathfrak{P}_2', \mathfrak{M}], \qquad \mathfrak{M} = M/M_1,$$

where $\mathfrak{M}$ is not divisible by $\mathfrak{P}_1$ or $\mathfrak{P}_2'$, it follows after multiplication by M_1

$$[A_1, M_1] = [A_2', M_1],$$

and A_1 and A_2' are directly semi-similar.

From Theorem 12 it follows, for example, that the decomposition is unique if (i) the anticenter is cyclic and (ii) when A_i belongs to the cyclic group of order p, no other A_j contains a cyclic group of order p in its principal series with respect to G. Various other special conditions may be derived from Theorem 12.

6. **The dual theory. Direct decompositions.** According to the general principle of duality which we have stated there must exist a theory dual to the preceding decomposition theory. Let us indicate briefly how this theory may be formulated.

A left-hand decomposition of a group G is a representation

$$\mathfrak{G} = G/E = [\mathfrak{A}, \mathfrak{B}]_l \tag{19}$$

of $\mathfrak{G}$ as the left-hand union of two quotients

$$\mathfrak{A} = G/A, \qquad \mathfrak{B} = G/B.$$

The relation (19) means, however,

$$(A, B) = E,$$

so that a left-hand decomposition of G is equivalent to the representation of the unit element E of G as the cross-cut of two relatively prime groups. One finds that G is indecomposable if and only if any two (normal) subgroups of G have a common subgroup. When the finite chain condition is satisfied for normal subgroups in G, this means that G shall have only one minimal normal subgroup.[7]

The representation of G as the left-hand union of left-hand indecomposable factors

$$\mathfrak{G} = [\mathfrak{A}_1, \cdots, \mathfrak{A}_n]_l, \qquad \mathfrak{A}_i = G/A_i$$

is equivalent to a representation

$$(A_1, \cdots, A_n) = E, \tag{20}$$

where A_i cannot be represented as the cross-cut of larger normal subgroups of G. One finds that when two such representations (20) exist both representations must have the same number of terms, the terms may be suitably

[7] Considerations of this nature can be found in the following papers by R. Remak, *Über die Darstellung der endlichen Gruppen als Untergruppen direkter Produkte*, Journal für Math., vol. 163(1930), pp. 1–44; *Über die erzeugenden invarianten Untergruppen der subdirekten Darstellungen endlicher Gruppen*, ibid., vol. 164(1931), pp. 197–242.

interchanged, and corresponding groups of the two representations have similar minimal normal subgroups. The representation is unique when G has no normal Abelian subgroups. When G has a (lower) cover, the number of terms in (20) is equal to the number of basis elements P_i of the cover and each A_i contains a normal subgroup similar to P_i.

In conclusion, one should also mention the properties of the direct decompositions. A very simple proof for the Schmidt-Remak theorem for groups in which the normal chains are finite can be obtained by a direct translation of the structure proof. This proof can also be extended to the case of certain infinite chains by the methods of a second paper on direct decomposition in Dedekind structures. For groups, however, these methods add very little to the general results obtained by Fitting[8] and Kořínek[9] and hence we shall omit a further discussion. Let us only mention the close connection of their method with the one applied in the second paper on direct decompositions in structures. In principle this method seems to go back to Krull.[10]

Chapter 3. Structural properties

1. **Distributive pairs of subgroups.** The general characterization of groups by means of their structural properties is an important problem of considerable difficulty. In the following we shall consider only certain questions connected with the distributive law in groups.

Let A and B be two subgroups of G. A subgroup T of $[A, B]$ shall be said to be *distributed* in $[A, B]$ when there exists a representation

$$T = [A_1, B_1], \tag{1}$$

where A_1 is a subgroup of A and B_1 a subgroup of B. Besides (1) there must then also exist a representation

$$T = [(A, T), (B, T)]. \tag{2}$$

This definition shows that if two subgroups T_1 and T_2 of $[A, B]$ are distributed, their union $[T_1, T_2]$ has the same property.

The distributive relation for three subgroups

$$(C, [A, B]) = [(C, A), (C, B)] \tag{3}$$

expresses that $(C, [A, B])$ is a distributed subgroup of $[A, B]$. The preceding remarks show that the study of the validity of the distributive relation (3) is equivalent to the study of the distributed subgroups of $[A, B]$. We shall say

[8] See for example H. Fitting, *Über die direkten Produktzerlegungen einer Gruppe in direkt unzerlegbare Faktoren*, Math. Zeitschrift, vol. 39(1935), pp. 16–40; or H. Zassenhaus, *Lehrbuch der Gruppentheorie*, pp. 78–82.

[9] V. Kořínek, *Sur la décomposition d'un groupe en produit direct de sousgroupes*, Časopis pro Pěstování matematiky a fysiky, vol. 66(1937), pp. 261–286.

[10] W. Krull, *Über verallgemeinerte endliche Abelsche Gruppen*, Math. Zeitschrift, vol. 23(1925), pp. 161–196. A similar method can be developed for arbitrary Dedekind structures: see O. Ore, *Direct decompositions*, this Journal, vol. 2(1936), pp. 581–596.

that the two groups A and B are *distributive* or form a *distributive pair* when the relation (3) holds for all C. This is equivalent to saying that every subgroup of $[A, B]$ is distributed.

We shall now consider the conditions for two groups A and B to be distributive. Let c be any element of $[A, B]$ not contained in A and B. Since the cyclic group $\{c\}$ must be representable in the form (2) when A and B are distributive, we must have relations

$$c^{n_A} = a, \qquad c^{n_B} = b, \tag{4}$$

where a is in A and b in B. The minimal exponents n_A and n_B for which (4) holds shall be called the orders of c with respect to A and B. According to (2) the cyclic groups $\{a\}$ and $\{b\}$ must generate $\{c\}$. Since a and b are powers of c, they are commutative and there must exist a representation

$$c = a^x \cdot b^y. \tag{5}$$

This fact establishes, first, that every element in $[A, B]$ has the form $c = a_1 \cdot b_1$, where a_1 belongs to A and b_1 to B. Hence we have proved

THEOREM 1. *If A and B form a distributive pair, the two groups are permutable.*

The relation (5) implies, secondly, the existence of integers x and y such that

$$x \cdot n_A + y \cdot n_B = 1,$$

and n_A and n_B are relatively prime.

THEOREM 2. *The necessary and sufficient condition that A and B form a distributive pair is that any element c in $[A, B]$ but not in A or B have finite orders n_A and n_B with respect to A and B such that $(n_A, n_B) = 1$.*

We have already shown that this condition is necessary. Conversely, if it is satisfied, every cyclic subgroup of $[A, B]$ is distributed and since every group is the union of its cyclic groups the theorem follows in general by a preceding remark.

It may happen that a distributive pair A, B contains subgroups A_1, B_1 such that

$$[A, B] = [A_1, B_1],$$

and such that A_1 and B_1 still is a distributive pair. We shall say that A_1 and B_1 have been obtained by *reduction* from A and B. Conversely, if A and B is a distributive pair, the groups $[A, D_B]$ and $[B, D_A]$ form a distributive pair when D_B is any subgroup of B and D_A any subgroup of A. This last pair shall be said to have been obtained by *extension* from A and B. We shall now show how a distributive pair may be obtained by extension from a particularly simple pair.

We first prove the following lemma:

Those elements $\bar{a}$ in A for which $b\bar{a}b^{-1}$ is contained in A for any b form a normal subgroup of $[A, B]$, when A and B are permutable groups.

It is obvious that the $\bar{a}$ form a group $\bar{A}$ for which

$$b\bar{A}b^{-1} = \bar{A}$$

for any b. To show that

$$a\bar{A}a^{-1} = \bar{A}$$

for any a, let us write $a' = a\bar{a}a^{-1}$. Then

$$ba'b^{-1} = bab^{-1} \cdot b\bar{a}b^{-1} \cdot (bab^{-1})^{-1}.$$

Here $b\bar{a}b^{-1} = \bar{a}_1$ is contained in $\bar{A}$. Furthermore, since A and B are permutable we can write

$$bab^{-1} = a_1 \cdot b_1 ,$$

and we find

$$ba'b^{-1} = (a_1b_1) \cdot \bar{a}_1 \cdot (a_1b_1)^{-1} = a_1 \cdot \bar{a}_2 \cdot a_1^{-1}.$$

Hence a belongs to A.

From now on we shall suppose that A and B are *finite* groups. We consider an element a in A which does not belong to $\bar{A}$ or to the cross-cut $D = (A, B)$. Then there exist elements b such that bab^{-1} does not belong to A or B. This implies according to Theorem 2 that there exist two relatively prime exponents $n_A^{(b)}$ and $n_B^{(b)}$ such that

$$(6) \qquad ba^{n_A^{(b)}}b^{-1} = a_1 , \qquad ba^{n_B^{(b)}}b^{-1} = b_1 .$$

The last relation (6) shows that $n_B^{(b)} = n_B$ is the order of a with respect to D. The exponent $n_A^{(b)}$ in the first relation (6) will usually depend on b. But since B is finite there exists a least common multiple m_A of all $n_A^{(b)}$, and m_A and n_B are relatively prime. It is obvious that a^{m_A} belongs to $\bar{A}$ and furthermore one can find integers x and y such that

$$a = a^{m_A \cdot x} \cdot a^{n_B \cdot y}.$$

This shows that

$$A = [\bar{A}, D],$$

and $\bar{A}$ and D form a distributive pair.

This result implies that the distributive pair A, B can be reduced to $\bar{A}$, B. A similar reduction on B gives

$$[A, B] = [\bar{A}, \bar{B}],$$

where $\bar{A}$, $\bar{B}$ is a distributive pair and $\bar{A}$ and $\bar{B}$ are normal in $[A, B]$.

When A and B are normal in $[A, B]$, the condition for distributivity is easily described. The orders of any element $c = a \cdot b$ with respect to A and B are seen to be equal to the orders $n_D^{(b)}$ and $n_D^{(a)}$ of b and a with respect to D. This shows that all orders of the elements of A with respect to D are relatively prime

to those of the elements of B with respect to D. This is only the case when the orders of the two quotient groups A/D and B/D are relatively prime.

We sum up the preceding results in

THEOREM 3. *Let A and B be finite normal subgroups of $[A, B]$. The necessary and sufficient condition that A and B be distributive is that the orders of the quotient groups $A/(A, B)$ and $B/(A, B)$ be relatively prime. Any distributive pair A_1 and B_1 in a finite group can be obtained by extension from a normal pair.*

2. **Distributive group structures.** Let us now consider the problem of finding all groups in which the structure of all subgroups is distributive. The result is

THEOREM 4. *The necessary and sufficient condition that the structure of all subgroups of a group G be distributive is that G be Abelian and have the property that any finite subset of elements generate a cyclic group.*

Proof. In this theorem we make no assumption of finiteness. We show first that the subgroup $M = \{a, b\}$ generated by two elements is cyclic when all subgroups are distributive. Let us suppose first that bab^{-1} is not contained in $\{a\}$. Then it follows as before that there exist relatively prime powers n and m such that

$$ba^nb^{-1} = a^{n_1}, \qquad a^m = b^{m_1}. \tag{7}$$

This gives a representation

$$a = a^{xn} \cdot a^{my}, \qquad xn + my = 1,$$

which shows that bab^{-1} must be a power of a. Similarly, one finds that aba^{-1} is a power of b. This means that $\{a\}$ and $\{b\}$ must be normal subgroups of M. Furthermore, since $\{a\}$ and $\{b\}$ are distributive, we can assume that the exponents m and m_1 in (7) are relatively prime. Now one must have a relation

$$a \cdot b = b \cdot a \cdot d, \tag{8}$$

where d is some element in the cross-cut of $\{a\}$ and $\{b\}$. Since d is a power of both a and b, it is permutable with both elements. From (8) one obtains by induction

$$a \cdot b^i = b^i \cdot a \cdot d^i, \qquad a^j \cdot b = b \cdot a^j \cdot d^j,$$

and for $i = m_1$ and $j = m$ this implies

$$d^m = d^{m_1} = 1,$$

or $d = 1$. Thus a and b are permutable. Furthermore, it is not difficult to see that one can choose the generating elements of A and B so that M is the cyclic group generated by ab.

The converse is easily proved. Let A and B be subgroups of G and a and b an element of each. Since the group $\{a, b\}$ is cyclic, it follows that the orders of a and b over the common subgroup of $\{a\}$ and $\{b\}$ must be finite and relatively

prime. Hence the conditions of Theorem 2 are satisfied and A and B are distributive.

For finite groups Theorem 4 says simply

THEOREM 5. *The only finite groups in which the structure of all subgroups is distributive are the cyclic groups.*

For the infinite groups with distributive structure we have found a special class of Abelian groups which are identical with the *ideal cyclic groups* introduced by Prüfer.

3. **Primitive elements.** Let G be a given group and let Σ be any substructure of the structure of all subgroups of G, i.e., when the set Σ contains the two subgroups A and B, it also contains $[A, B]$ and (A, B). We shall say that an element a of a group A belonging to Σ is a *primitive element* of A in Σ, when a belongs to A, but not to any subgroup of A contained in Σ.

We shall say that a group A is *indecomposable in* Σ, when it is not the union of two proper subgroups both contained in Σ. If we suppose that all ascending chains in Σ are finite, we find as before, in the normal case, that the necessary and sufficient condition for A to be indecomposable is that it contain but a single maximal subgroup B contained in Σ. This makes it obvious that a group A which is indecomposable in Σ must contain primitive elements with respect to Σ, since any element of A not contained in B must be primitive.

We shall now prove the following

THEOREM 6. *Let Σ be a substructure of the structure of all subgroups of a given group G. We suppose that Σ satisfies the Dedekind axiom and that all chains are finite. Then any subgroup A in Σ whose representation as the union of indecomposable groups in Σ*

$$A = [A_1, \cdots, A_r] \tag{9}$$

is unique must contain primitive elements with respect to Σ.

Proof. Let B_i be the maximal group of A_i with respect to Σ. Furthermore, let us write

$$\bar{A}_i = [A_1, \cdots, A_{i-1}, A_{i+1}, \cdots, A_r].$$

It then follows from the Dedekind axiom that the groups

$$M_i = [B_i, \bar{A}_i]$$

are maximal subgroups of A in Σ. We shall show further that the uniqueness of the decomposition (9) implies that the groups M_i are the only maximal groups.

Let us denote by K some maximal group of A in Σ. Since K does not contain all A_i, let us suppose that it does not contain A_1. We shall show that $K = M_1$. Since

$$A = [A_1, K],$$

it follows that if K has the decomposition

$$K = [K_1, \cdots, K_s]$$

into indecomposable groups in Σ, then all A_i $(i \geqq 2)$ must occur among the K_i. This means that K must contain $\bar{A}_1$. It must also contain B_1 because otherwise

$$A = [K, B_1]$$

would give a new decomposition of A.

The existence of primitive elements in A now follows simply. Since A_i is not contained in M_i, let α_i be an element of A_i not in M_i. Then the product

$$\gamma = \alpha_1 \cdots \alpha_n$$

is a primitive element because it cannot be contained in any maximal group M_i.

From Theorem 6 we deduce

THEOREM 7. *Let Σ be a distributive substructure of finite length in the structure of all subgroups of a group G. Then every group belonging to Σ contains primitive elements with respect to Σ.*

The proof follows from the observation that in a distributive structure any decomposition must be unique.

YALE UNIVERSITY.

SUMMABILITY OF CONJUGATE FOURIER SERIES

By A. H. Smith

1. **Introduction.** We assume that the function $f(x)$ is integrable in the sense of Lebesgue over $(0, 2\pi)$ and satisfies the periodicity condition $f(x + 2\pi) = f(x)$; then the series

$$\sum_{\nu=1}^{\infty} (a_\nu \sin \nu x - b_\nu \cos \nu x), \tag{1}$$

where a_ν, b_ν are the Fourier coefficients, is defined to be the conjugate Fourier series generated by $f(x)$. Let us also define, whenever the limits on the right side exist,

$$g(x) \equiv \lim_{\eta \to 0} -\frac{1}{\pi}\int_\eta^\infty \frac{\psi(t)}{t}\, dt, \tag{2}$$

and

$$G(x) \equiv \lim_{\eta \to 0} -\frac{1}{\pi}\int_\eta^\infty \frac{\Psi(t)}{t^2}\, dt, \tag{3}$$

where

$$\psi(t) \equiv f(x + t) - f(x - t); \qquad \Psi(t) \equiv \int_0^t \psi(u)\, du. \tag{4}$$

Paley[1] has proved that the series (1) is summable by Cesàro means of order α for $\alpha > 1$ to $g(x)$, whenever $g(x)$ exists. From two lemmas due to Prasad[2] it is easily shown that the existence of $G(x)$ is equivalent to the existence of $g(x)$ with the added condition that $\Psi(t) = o(t)$ as $t \to 0$. Prasad generalized Paley's result by proving that whenever $G(x)$ exists the series (1) is summable by Cesàro means of order α for $\alpha > 1$ to $G(x)$, provided that either

(i) $\Psi(t) = O(t)$ as $t \to 0$, or

(ii) $\displaystyle\int_0^t \frac{|\Psi(t)|}{t}\, dt = o(t)$ as $t \to 0$.

Received September 13, 1937.

[1] R. E. A. C. Paley, *On the Cesàro summability of Fourier series and allied series*, Proceedings of the Cambridge Philosophical Society, vol. 26(1930), pp. 173–203.

[2] B. N. Prasad, *Contribution à l'étude de la série conjuguée d'une série de Fourier*, Journal de Mathématiques Pures et Appliquées, (9), vol. 11(1932), pp. 153–205 (Lemmas 1 and 2). Prasad shows that whenever $\lim_{\eta\to 0}\int_\eta^\delta \psi(t)t^{-1}\, dt$ exists, then $\lim_{\eta\to 0}\int_\eta^\delta \psi(t)t^{-2}\, dt$ exists, but when the latter exists the necessary and sufficient condition for the existence of the former is that $\Psi(t) = o(t)$ as $t \to 0$.

The purpose of this paper is to derive more exact information concerning the summability of the series (1) under the hypothesis just stated. We use the summation method of Bosanquet-Linfoot,[3] which is weaker than Cesàro's. Our results are given in Theorems 1 and 2. In §4 we give an example of a function $\Psi(t)$ which satisfies the conditions:

(i) $\Psi(t) = O(t)$ as $t \to 0$;

(ii) $\displaystyle\int_0^t \frac{|\Psi(t)|}{t}\, dt = o(t)$ as $t \to 0$;

(iii) $G(x)$ exists.

However, $\Psi(t) = o(t)$ as $t \to 0$ is not satisfied, and hence $g(x)$ does not exist.

2. Definitions and lemmas. The following functions will be used throughout the paper:

$$H_{k,\alpha,\beta}(1-u) \equiv Bu^k(1-u)^{\alpha-1}\log^{-\beta}\left(\frac{C}{1-u}\right), \tag{5}$$

where $B = (\log C)^\beta$, for $k \geqq 0$, $\alpha \geqq 1$, $\beta \geqq 0$;

$$\left.\begin{matrix} Q_{k,\alpha,\beta}(t) \\ \bar{Q}_{k,\alpha,\beta}(t) \end{matrix}\right\} \equiv \int_0^1 H_{k,\alpha,\beta}(1-u) \left.\begin{matrix} \cos tu \\ \sin tu \end{matrix}\right\} du; \tag{6}$$

$$\left.\begin{matrix} \lambda_{\alpha,\beta}(n,t) \\ \bar{\lambda}_{\alpha,\beta}(n,t) \end{matrix}\right\} \equiv \frac{1}{\pi}\int_0^n H_{0,\alpha,\beta}(1-\nu/n) \left.\begin{matrix} \cos \nu t \\ \sin \nu t \end{matrix}\right\} d\nu. \tag{7}$$

The functions $\lambda^{(k)}_{\alpha,\beta}(n, t)$ and $\bar{\lambda}^{(k)}_{\alpha,\beta}(n, t)$ denote the k-th derivatives with respect to t of $\lambda_{\alpha,\beta}(n, t)$ and $\bar{\lambda}_{\alpha,\beta}(n, t)$, respectively. It is readily seen that $\lambda^{(k)}_{\alpha,\beta}(n, t)$ and $\bar{\lambda}^{(k)}_{\alpha,\beta}(n, t)$ can be expressed in terms of either $Q_{k,\alpha,\beta}(nt)$ or $\bar{Q}_{k,\alpha,\beta}(nt)$.

We have need also of the following lemmas, the first two of which are stated without proof.[4]

LEMMA 1. *For $k \geqq 0$, $\alpha \geqq 1$, $\beta \geqq 0$, the functions $Q_{k,\alpha,\beta}(t)$ and $\bar{Q}_{k,\alpha,\beta}(t)$ are bounded in $(0, \infty)$, and for large values of t,*

$$Q_{k,\alpha,\beta}(t) + i\bar{Q}_{k,\alpha,\beta}(t) = \frac{i^{k+1}k!}{t^{k+1}} + \frac{i^{k+2}c_{k+2}}{t^{k+2}} + O\left(\frac{1}{t^{k+3}}\right) + O\left(\frac{1}{t^\alpha \log^\beta t}\right),$$

where c_{k+2} is a constant.

[3] L. S. Bosanquet and E. H. Linfoot, *Generalized means and the summability of Fourier series*, Quarterly Journal of Mathematics, (2), vol. 2(1931), pp. 207–229. The series Σa_ν is said to be summable (α, β) to S, where either $\alpha > 0$, or $\alpha = 0, \beta \geqq 0$, if

$$\sum_{\nu<n}\left[B(1-\nu/n)^\alpha \log^{-\beta}\left(\frac{C}{1-\nu/n}\right)a_\nu\right] \to S \quad \text{as } n \to \infty$$

for C sufficiently large, where $B = (\log C)^\beta$. They have shown that it is equivalent to say "for every $C > 1$". (Theorem 3.2.)

[4] Lemma 1 is Lemma 2.1 and the proof of Lemma 2 is analogous to that of Lemma 2.2 of the following paper: A. H. Smith, *On the summability of derived series of the Fourier-Lebesgue type*, Quarterly Journal of Mathematics, (2), vol. 4(1933), pp. 93–106.

Lemma 2. *The function $\bar{\lambda}^{(k)}_{\alpha,\beta}(n, t)$ is bounded in $(0, \infty)$ for fixed n, where $k \geqq 0$, $\alpha \geqq 1$, $\beta \geqq 0$; and for large values of t, where $k \geqq 0$, and $\beta \geqq 0$,*

$$\bar{\lambda}^{(k)}_{\alpha,\beta}(n, t) = \begin{cases} \dfrac{(-1)^k k!}{\pi t^{k+1}} + O(n^{-\delta} t^{-(k+1+\delta)} \log^{-\beta}(nt)) & (\alpha = k + 1 + \delta, 0 \leqq \delta < 2), \\ \dfrac{(-1)^k k!}{\pi t^{k+1}} + O(n^{-2} t^{-(k+3)}) & (\alpha \geqq k + 3). \end{cases}$$

Lemma 3. *The function $\bar{\lambda}_{\alpha,\beta}(n, t)$ can be expressed as follows:*

$$\bar{\lambda}_{\alpha,\beta}(n, t) = \frac{1}{\pi t} - \frac{\alpha - 1}{nt} \lambda_{\alpha-1,\beta}(n, t) - \frac{\beta}{nt} \lambda_{\alpha-1,\beta+1}(n, t).$$

Proof. From (7)

$$\bar{\lambda}_{\alpha,\beta}(n, t) = \frac{n}{\pi} \int_0^1 H_{0,\alpha,\beta}(1 - u) \sin unt\, du.$$

The lemma is obtained upon integration by parts.

Lemma 4. *For $\beta > 1$ the function $Q_{0,1,\beta}(t)$ is of bounded variation over $(0, \infty)$.*

Proof. It is easily seen that

$$Q^{(1)}_{0,1,\beta}(t) = -\bar{Q}_{1,1,\beta}(t).$$

Therefore, by Lemma 1, $Q^{(1)}_{0,1,\beta}(t)$ is bounded over $(0, \infty)$ and is $O(t^{-1} \log^{-\beta} t)$ for large values of t. Hence

$$\int_0^\infty | Q^{(1)}_{0,1,\beta}(t) |\, dt$$

exists for $\beta > 1$ and thus the lemma is proved.

Lemma 5. *For $\beta > 1$ the function $t\bar{Q}_{1,1,\beta}(t)$ is of bounded variation over any fixed interval.*

Proof. As in Lemma 4 we show that $\bar{Q}_{1,1,\beta}(t)$ is of bounded variation over $(0, \infty)$ for $\beta > 1$. The lemma then follows.

Lemma 6. *The expression*

$$-\int^\infty \psi(t)\bar{\lambda}_{2,\beta}(n, t)\, dt$$

is the $[n]$-th[5] mean of order (α, β), for $\alpha = 1$, $\beta \geqq 0$, of the conjugate Fourier series.

The proof of this lemma is similar to that of one in an earlier paper.[6]

[5] Where $[n]$ denotes the largest integer $\leqq n$.

[6] A. H. Smith, *On the summability of derived conjugate series of the Fourier-Lebesgue type*, Bulletin of the American Mathematical Society, vol. 40(1934), pp. 406–412. In particular see Lemma 4, p. 409 and §5, p. 412. Note that the condition "that $f(x)$ be finite" of Lemma 4 is needed only in the case of the derived conjugate series.

3. Summability theorems. We now proceed to develop our main results. From Lemma 6 we obtain

$$-\int_0^\infty \psi(t)\bar{\lambda}_{2,\beta}(n, t)\, dt = -[\Psi(t)\bar{\lambda}_{2,\beta}(n, t)]_0^\infty + \int_0^\infty \Psi(t)\bar{\lambda}_{2,\beta}^{(1)}(n, t)\, dt.$$

The integrated term vanishes at the lower limit; it also vanishes at the upper limit, since $\Psi(t)$ is periodic and for large values of t (Lemma 2)

$$\bar{\lambda}_{2,\beta}(n, t) = \frac{1}{\pi t} + O(n^{-1}t^{-2}\log^{-\beta} nt).$$

Thus the $[n]$-th mean of order $(1, \beta)$, for $\beta \geqq 0$, of the conjugate Fourier series can be expressed as

$$I \equiv \int_0^\infty \Psi(t)\bar{\lambda}_{2,\beta}^{(1)}(n, t)\, dt. \tag{8}$$

THEOREM 1. *If the function $f(x)$ is integrable in the sense of Lebesgue and has period 2π, the conjugate series is summable (α, β) for $\alpha = 1$, $\beta > 1$ to $G(x)$ provided that*

(i) *$G(x)$ exists;*

(ii) *$\Psi(t) = O(t)$ as $t \to 0$.*

Proof. Assume that at the point x the conditions of the theorem are satisfied; let K denote a positive numerical constant. Choose A so that

$$\Psi(t) < Kt \qquad \text{for } 0 \leqq t \leqq A; \tag{9}$$

then choose n so that $nA > e$ and define, where $\delta > 0$,

$$\begin{cases} I_1 \equiv \displaystyle\int_0^{en^{\delta-1}} \Psi(t)\bar{\lambda}_{2,\beta}^{(1)}(n, t)\, dt, \\ I_2 \equiv \displaystyle\int_{en^{\delta-1}}^{A} \Psi(t)\bar{\lambda}_{2,\beta}^{(1)}(n, t)\, dt, \\ I_3 \equiv \displaystyle\int_A^\infty \Psi(t)\bar{\lambda}_{2,\beta}^{(1)}(n, t)\, dt. \end{cases} \tag{10}$$

We proceed to investigate I_1. Since by Lemma 3

$$\bar{\lambda}_{2,\beta}^{(1)}(n, t) = -\frac{1}{\pi t^2} + \frac{\lambda_{1,\beta}(n, t)}{nt^2} - \frac{\lambda_{1,\beta}^{(1)}(n, t)}{nt} + \frac{\beta\lambda_{1,\beta+1}(n, t)}{nt^2} - \frac{\beta\lambda_{1,\beta+1}^{(1)}(n, t)}{nt},$$

we divide I_1 into the respective integrals $I_{1.1}$, $I_{1.2}$, $I_{1.3}$, $I_{1.4}$ and $I_{1.5}$. Because of the existence of $G(x)$ the integral

$$I_{1.1} \equiv -\frac{1}{\pi}\int_0^{en^{\delta-1}} \frac{\Psi(t)}{t^2}\, dt = o(1) \qquad \text{as } n \to \infty.$$

Investigating $I_{1.2}$, we substitute

$$\lambda_{1,\beta}(n, t) = nQ_{0,1,\beta}(nt)$$

((6) and (7)) and obtain

$$I_{1.2} = \frac{1}{n}\int_0^{en^\delta} \Psi\left(\frac{t}{n}\right)\left(\frac{t}{n}\right)^{-2} Q_{0,1,\beta}(t)\,dt.$$

By Lemma 4, $Q_{0,1,\beta}(t)$, for $\beta > 1$, can be replaced by $P(t) - N(t)$, where $P(t)$ and $N(t)$ are positive monotone decreasing bounded functions. Now

$$\frac{1}{n}\int_0^{en^\delta} \Psi\left(\frac{t}{n}\right)\left(\frac{t}{n}\right)^{-2} P(t)\,dt = P(0)\,\frac{1}{n}\int_0^{\xi} \Psi\left(\frac{t}{n}\right)\left(\frac{t}{n}\right)^{-2} dt, \qquad 0 \leqq \xi < en^\delta,$$

$$= P(0)o(1) = o(1) \qquad \text{as } n \to \infty,$$

since $G(x)$ exists. Thus $I_{1.2} = o(1)$ as $n \to \infty$, for $\beta > 1$. Similarly $I_{1.4} = o(1)$ as $n \to \infty$, for $\beta > 0$. Consider next $I_{1.3}$. Divide the interval $(0, en^{\delta-1})$ into $(0, ln^{-1})$ and $(ln^{-1}, en^{\delta-1})$, where we choose l so that $\log^{1-\beta} l < \epsilon$ for a preassigned ϵ, then n so that $ln^{-1} < en^{\delta-1}$. Since

$$\lambda_{1,\beta}^{(1)}(n, t) = n^2\bar{Q}_{1,1,\beta}(nt), \tag{11}$$

it follows that

$$\int_0^{ln^{-1}} \frac{\Psi(t)}{nt}\,\lambda_{1,\beta}^{(1)}(n, t)\,dt = \frac{1}{n}\int_0^{l} \Psi\left(\frac{t}{n}\right)\left(\frac{t}{n}\right)^{-2} t\bar{Q}_{1,1,\beta}(t)\,dt.$$

Because, for $\beta > 1$, $t\bar{Q}_{1,1,\beta}(t)$ is of bounded variation over any fixed interval, it may be replaced by $P(t) - N(t)$, where $P(t)$ and $N(t)$ are positive monotone decreasing bounded functions of t. Moreover,

$$\frac{1}{n}\int_0^{l} \Psi\left(\frac{t}{n}\right)\left(\frac{t}{n}\right)^{-2} P(t)\,dt = P(0)o(1) = o(1) \qquad \text{as } n \to \infty, \text{ for } \beta > 1,$$

since $G(x)$ exists. The argument may be repeated for $N(t)$. From condition (ii) of the theorem and (11)

$$\int_{ln^{-1}}^{en^{\delta-1}} \frac{\Psi(t)}{nt}\,\lambda_{1,\beta}^{(1)}(n, t)\,dt = O\left(\int_l^{en^\delta} \bar{Q}_{1,1,\beta}(t)\,dt\right).$$

If we substitute, from Lemma 1, $\bar{Q}_{1,1,\beta}(t) = O(t^{-1}\log^{-\beta} t)$ and integrate, this integral becomes arbitrarily small with ϵ for $\beta > 1$ on account of the choice of l. Thus we have shown that $I_{1.3} = o(1)$ as $n \to \infty$. The integral $I_{1.5}$ may be treated in a like manner. Hence we have shown that for $\beta > 1$

$$I_1 = o(1) \qquad \text{as } n \to \infty. \tag{12}$$

We next investigate I_2. By Lemma 2

$$\bar{\lambda}_{2,\beta}^{(1)}(n, t) = -\frac{1}{\pi t^2} + O\left(\frac{1}{t^2 \log^\beta nt}\right). \tag{13}$$

Set $I_2 = I_{2.1} + I_{2.2}$, where

$$I_{2.1} = -\frac{1}{\pi}\int_{en^{\delta-1}}^{A} \frac{\Psi(t)}{t^2}\,dt; \qquad I_{2.2} = O\left(\int_{en^{\delta-1}}^{A} \frac{|\,\Psi(t)\,|}{t^2\log^\beta nt}\,dt\right).$$

Now by condition (ii) of the theorem

$$I_{2.2} = O\left(\int_{en^{\delta-1}}^{A} \frac{1}{t \log^{\beta} nt}\, dt\right) = o(1) \qquad \text{as } n \to \infty, \text{ for } \beta > 1.$$

Thus

$$(14) \qquad I_2 \to I_{2.1} \qquad \text{as } n \to \infty.$$

Finally consider I_3. Using (3), we set $I_3 = I_{3.1} + I_{3.2}$, where $I_{3.1}$ and $I_{3.2}$ are defined similarly to $I_{2.1}$ and $I_{2.2}$. Now choose q so that $2(q-1)\pi \leqq A < 2q\pi$; then, since $\Psi(t)$ is periodic, K being a constant,

$$|I_{3.2}| \leqq K\left[\int_0^{2\pi} |\Psi(t)|\, dt \left\{\frac{1}{A^2 \log^{\beta} nA} + \frac{1}{(2\pi)^2} \sum_{\nu=q}^{\infty} \frac{1}{\nu^2 \log^{\beta} 2n\pi\nu}\right\}\right].$$

Thus $I_{3.2} \to 0$ as $n \to \infty$, for $\beta > 0$. Hence under these conditions

$$(15) \qquad I_3 \to I_{3.1}.$$

Combining (10), (12), (14) and (15), we have, for $\beta > 1$,

$$\lim_{n\to\infty} I = \lim_{n\to\infty} -\frac{1}{\pi}\int_{en^{\delta-1}}^{\infty} \frac{\Psi(t)}{t^2}\, dt = \lim_{\eta\to 0} -\frac{1}{\pi}\int_{\eta}^{\infty} \frac{\Psi(t)}{t^2}\, dt,$$

provided the latter limit exists. This completes the proof of the theorem.

THEOREM 2. *If the function $f(x)$ is integrable in the sense of Lebesgue and has period 2π, the conjugate series is summable (α, β) for $\alpha = 1$, $\beta > 1$ to $G(x)$ provided that*

(i) *$G(x)$ exists;*

(ii) $\displaystyle\int_0^t \frac{|\Psi(t)|}{t}\, dt = o(t)$ *as $t \to 0$.*

Proof. Assuming that at the point x the conditions of the theorem are satisfied, choose ϵ, then A so that

$$\chi(t) = \int_0^t \frac{|\Psi(t)|}{t}\, dt < \epsilon t \qquad \text{for } 0 \leqq t \leqq A.$$

Divide the interval $(0, \infty)$ into $(0, en^{-1})$, (en^{-1}, A) and (A, ∞). Set $I = J_1 + J_2 + J_3$, where J_1, J_2, J_3 are defined in a manner analogous to I_1, I_2, I_3. It can be shown, as in the case of I_1, that $J_1 \to 0$ as $n \to \infty$, for $\beta > 1$.[7]

Next consider J_2. Using (13), we set $J_2 = J_{2.1} + J_{2.2}$, where

$$J_{2.1} = -\frac{1}{\pi}\int_{en^{-1}}^{A} \frac{\Psi(t)}{t^2}\, dt, \qquad J_{2.2} = O\left(\int_{en^{-1}}^{A} \frac{|\Psi(t)|}{t^2 \log^{\beta} nt}\, dt\right).$$

[7] The only change except for the upper limit is in the case of $J_{1.3}$ and $J_{1.5}$, which are treated as the first integral of $I_{1.3}$.

Now

$$
(16)\qquad \int_{en^{-1}}^{A} \frac{\Psi(t)}{t}\cdot\frac{1}{t\log^{\beta} nt}\,dt = \left[\frac{\chi(t)}{t\log^{\beta} nt}\right]_{en^{-1}}^{A} + \int_{en^{-1}}^{A} \chi(t)t^{-2}\log^{-\beta} nt\,dt + \beta\int_{en^{-1}}^{A} \chi(t)t^{-2}\log^{-(\beta+1)} nt\,dt.
$$

When $n \to \infty$ the integrated term vanishes at the upper limit for $\beta > 1$ and at the lower limit on account of condition (ii) of the theorem. But

$$
\left|\int_{en^{-1}}^{A} \chi(t)t^{-2}\log^{-\beta} nt\,dt\right| \leqq \epsilon\int_{en^{-1}}^{A} t^{-1}\log^{-\beta} nt\,dt,
$$

and for $\beta > 1$, this term is arbitrarily small with ϵ. Since the second integral on the right of (16) can be treated likewise, it follows that, for $\beta > 1$, $J_2 \to J_{2\cdot1}$ as $n \to \infty$.

The integral J_3 is similar to I_3. Hence the theorem has been demonstrated.

4. Prasad gave the following example[8]

$$
\Psi(t) = \begin{cases} \dfrac{1}{n^2}\sin^2\left\{n^4\left(t - \dfrac{1}{n^2}\right)\pi\right\}, & \dfrac{1}{n^2} \leqq t \leqq \dfrac{1}{n^2} + \dfrac{1}{n^4} \qquad (n = 1, 2, \cdots), \\ 0 & \text{elsewhere,} \end{cases}
$$

and showed that the function defined had these properties:

(i) $\Psi(t) = O(t)$ as $t \to 0$;

(ii) $G(x)$ exists;

(iii) $\Psi(t) = o(t)$ as $t \to 0$ is not satisfied.

Hence the integral $g(x)$ does not exist.

We wish to show that

$$
(17)\qquad \int_0^t \frac{|\Psi(t)|}{t}\,dt = o(t) \quad \text{as } t \to 0.
$$

Now for $(m + 1)^{-2} + (m + 1)^{-4} < t \leqq m^{-2} + m^{-4}$,

$$
\int_0^t \frac{|\Psi(t)|}{t}\,dt \leqq \sum_{n=m}^{\infty} \frac{1}{n^2}\int_{n^{-2}}^{n^{-2}+n^{-4}} \frac{\sin^2[n^4(t - n^{-2})\pi]}{t}\,dt < \sum_{n=m}^{\infty} \frac{1}{2n^4}.
$$

Thus

$$
\frac{1}{t}\int_0^t \frac{|\Psi(t)|}{t}\,dt < \frac{\frac{1}{2}\sum_{n=m}^{\infty} n^{-4}}{(m+1)^{-2}} < \epsilon \qquad \text{for } n \text{ sufficiently large.}
$$

Hence condition (17) is satisfied.

PURDUE UNIVERSITY.

[8] B. N. Prasad, loc. cit., Chapter 4, §6.

A DETERMINATION OF THE AUTOMORPHISMS OF CERTAIN ALGEBRAIC FIELDS

By Caroline A. Lester

1. **Introduction.** The problem of determining all equations with rational coefficients having a given Galois group has been solved only in a few cases. All normal cubic and quartic equations having a prescribed group have been explicitly determined [8][1] as equations whose coefficients are rational in certain parameters. The possibility of so expressing all equations of higher degree having a prescribed group was discussed by E. Noether [7]. Less explicit determinations of normal equations of degrees five [5] and eight [4] have been made.

The companion problem of determining explicitly the automorphisms of a given normal field in terms of the coefficients of the defining equation has met with less success. It has been solved for the cyclic cubic [9]. Generating automorphisms for cyclic quartic and octic fields were found by Albert [1] and for cyclic quintic fields by Hull [5], but not explicitly in terms of the coefficients.

In the present paper, all automorphisms are explicitly obtained by purely algebraic methods for the cyclic cubic, quartic, and sextic, the quartic with the four-group, the sextic with the symmetric group, and the octics with the Abelian groups of types (2, 2, 2) and (2, 4). The determination of the parametric representations of the most general equations defining these fields was an integral part of the determination of the automorphisms, and while for $n = 3$ and $n = 4$ these results merely confirm known facts, the purely rational method of their attainment should not be without interest.

The computation in the cases $n = 6$ and $n = 8$ was brought within practicable limits by a free use of matric theory; in particular, of a theorem of Williamson [10]. This was used with particular success in the concluding section to obtain from the results of Albert a one-parameter family of cyclic octics over the rational field together with their automorphisms.

The writer is indebted to Professor C. C. MacDuffee, under whose direction this investigation was carried out.

2. **Cyclic cubics.** If the reduced cubic

$$f(x) = x^3 + px + q = 0 \tag{1}$$

is normal over the rational field F, it has the roots

$$\alpha, \qquad \alpha' = \theta_1(\alpha), \qquad \alpha'' = \theta_2(\alpha),$$

Received October 7, 1937; presented to the American Mathematical Society, October 30, 1937.

[1] The numbers in brackets refer to the bibliography at the end.

where θ_1 and θ_2 are polynomials of degrees $\leqq 2$. In $F(\alpha)$,

$$f(x) = (x - \alpha)(x^2 + \alpha x + p + \alpha^2) = 0.$$

Then $\theta_1(\alpha) = a\alpha^2 + b\alpha + c$ is a zero of the second factor. Upon substituting θ_1 for x in this factor and reducing modulo $f(\alpha)$, we obtain a polynomial, quadratic in α, which must vanish identically. This gives

$$\begin{aligned} b^2 + b + 1 + 2ac - a^2p &= 0, \\ 2bc + c - a^2q - 2abp - ap &= 0, \\ c^2 + p - 2abq - aq &= 0. \end{aligned} \tag{2}$$

The eliminant of this system is the determinant of the coefficients of 1, p, q. Thus

$$a^2(2b^2 + 2b + ac + 1)(2b^2 + 2b + ac + 2) = 0. \tag{3}$$

If $a = 0$, then $b^2 + b + 1 = 0$, so that b is not rational, and this case does not lead to a normal cubic over the rational field.

In case $2b^2 + 2b + ac + 1 = 0$ and $a \neq 0$, we find from (2),

$$p = -(3b^2 + 3b + 1)/a^2, \qquad q = (2b + 1)(b^2 + b)/a^2,$$

whence

$$a^3f(x) = (ax - b)(ax - b - 1)(ax + 2b + 1) = 0.$$

This case also leads to no normal cubic.

Now suppose $2b^2 + 2b + ac + 2 = 0$, $a \neq 0$. Then

$$(2b + 1)^2 = -2ac - 3 \tag{4}$$

and from (2)

$$p = 3c/2a, \qquad q = -c(2b + 1)/2a^2 \tag{5}$$

so that

$$2a^2f(x) = 2a^2x^3 + 3acx - c(2b + 1) = 0. \tag{6}$$

From (5),

$$c = 2ap/3, \qquad 3aq = -p(2b + 1).$$

Square this last relation and use (4), obtaining

$$9a^2q^2 = -p^2(2ac + 3).$$

But $2ac + 3 = (4a^2p + 9)/3$, so that

$$9p^2 = a^2(-4p^3 - 27q^2) = a^2d,$$

where d is the discriminant of (1). Then

$$a = \pm \frac{3p}{d^{\frac{1}{2}}}, \qquad c = \pm \frac{2p^2}{d^{\frac{1}{2}}}, \qquad b = \frac{\mp 9q - d^{\frac{1}{2}}}{2d^{\frac{1}{2}}}. \tag{7}$$

Hence the roots of (1) can be written [9]

$$\begin{aligned} \alpha, \quad \theta_1(\alpha) &= \frac{6p\alpha^2 - (9q + d^{\frac{1}{2}})\alpha + 4p^2}{2d^{\frac{1}{2}}}, \\ \theta_2(\alpha) &= \frac{-6p\alpha^2 + (9q - d^{\frac{1}{2}})\alpha - 4p^2}{2d^{\frac{1}{2}}}. \end{aligned} \tag{8}$$

These are rational functions of α if and only if d is a square in F. Hence (1) is cyclic if and only if it is irreducible and its discriminant is a rational square, a well-known result.

The birational transformation $y = 2ax$ puts (6) into the form

$$y^3 + 6acy - 4ac(2b + 1) = 0.$$

Set $ac = e$. We have

THEOREM 1. *If the equation*

$$f(x) = x^3 + 6ex - 4e(2b + 1) = 0, \tag{9}$$

where e and b assume all rational values such that $(2b + 1)^2 = -2e - 3$, *is irreducible, it is cyclic and defines all cyclic cubic fields.*

By (8), the automorphisms of (9) are

$$\alpha, \qquad \theta_1(\alpha) = \tfrac{1}{2}\alpha^2 + b\alpha + 2e, \qquad \theta_2(\alpha) = -\tfrac{1}{2}\alpha^2 - (b + 1)\alpha - 2e. \tag{10}$$

It is readily verified that they generate the cyclic group of order 3.

3. Normal quartics. A normal equation of degree 4 is either cyclic or has the 4-group as its Galois group. In either case the group is imprimitive, and the quartic can be put [11] into the form

$$f(x) = x^4 + px^2 + q = 0, \tag{11}$$

where p and q are rational. If $f(\alpha) = 0$, then $f(-\alpha) = 0$ and, since the automorphisms form an Abelian group, every automorphism $\theta(\alpha)$ must be an odd function of α.

The conjugates of α can be written

$$\alpha, \qquad \alpha' = \theta(\alpha) = a\alpha^3 + b\alpha, \qquad \alpha'' = -\alpha, \qquad \alpha''' = -\theta(\alpha), \tag{12}$$

where a and b are rational. Write

$$f(x) = (x - \alpha)(x + \alpha)(x^2 + p + \alpha^2) = 0.$$

Then $\theta(\alpha)$ is a zero of the last factor. Set $x = a\alpha^3 + b\alpha$ and reduce modulo $f(\alpha)$. We obtain a linear polynomial in α^2 which must vanish identically, i.e.,

$$(13)\qquad b^2 - 2abp + a^2p^2 - a^2q + 1 = 0, \qquad 2abq - a^2pq - p = 0.$$

If $a = 0$, $b^2 + 1 = 0$, so that b is not rational. Hence we may assume $a \neq 0$. Eliminate b from (13) and obtain

$$a^4(p^2q^2 - 4q^3) - 2a^2(p^2q - 2q^2) + p^2 = 0.$$

Formal solution gives

$$a = \pm p(p^2q - 4q^2)^{-\frac{1}{2}}, \qquad a = \pm q^{-\frac{1}{2}}.$$

Case I, $a = \pm p(p^2q - 4q^2)^{-\frac{1}{2}}$. If $p^2q - 4q^2 = 0$, either $q = 0$ or $p^2 = 4q$, and in either case (11) is reducible. If a is to be rational, $p^2q - 4q^2$ is a rational square $\neq 0$. In that case, (12) gives

$$(14)\quad \alpha, \qquad \alpha' = \theta(\alpha) = \frac{p\alpha^3 + (p^2 - 2q)\alpha}{(p^2q - 4q^2)^{\frac{1}{2}}}, \qquad \alpha'' = -\alpha, \qquad \alpha''' = -\theta(\alpha).$$

It is easily verified that these functions give a representation of the cyclic group of order 4.

Case II, $a = \pm q^{-\frac{1}{2}}$. We must have q a rational square. The two cases are mutually exclusive, for if $p^2q - 4q^2$ and q are both squares, $p^2 - 4q$ is a square and (11) is reducible. Now (12) becomes

$$(15)\qquad \alpha' = \theta(\alpha) = (\alpha^3 + p\alpha)q^{-\frac{1}{2}}, \qquad \alpha'' = -\alpha, \qquad \alpha''' = -\theta(\alpha).$$

These functions represent the 4-group.

Hence we have the known result [3]: the irreducible quartic (11) is cyclic if and only if $p^2q - 4q^2$ is a rational square, and is normal with the 4-group as its Galois group if and only if q is a rational square. All normal quartic fields are thus defined.

4. Cyclic quartics. From (14),

$$a^2 = \frac{p^2}{p^2q - 4q^2}, \qquad b^2 = \frac{(p^2 - 2q)^2}{p^2q - 4q^2},$$

whence

$$p = \frac{3b \pm (b^2 - 8)^{\frac{1}{2}}}{2a}, \qquad q = \frac{b^2 - 2 \pm b(b^2 - 8)^{\frac{1}{2}}}{2a^2}.$$

For p and q to be rational, $b^2 - 8$ must be a perfect square. Let $b = (2k^2 + 1)/k$, $k = [b \pm (b^2 - 8)^{\frac{1}{2}}]/4$. Then

$$p = \frac{4k_1^2 + 1}{ak_1}, \qquad q = \frac{4k_1^2 + 1}{a^2}, \qquad \text{or} \qquad p = \frac{2(k_2^2 + 1)}{ak_2}, \qquad q = \frac{k_2^2 + 1}{a^2k_2^2}.$$

The two sets of values for p and q yield the same result upon substituting $k_2 = 1/2k_1$. Hence (11) becomes

$$f(x) = x^4 + \frac{4k^2+1}{ak}x^2 + \frac{4k^2+1}{a^2} = 0.$$

Letting $y = akx$, we have

$$y^4 + ak(4k^2+1)y^2 + a^2k^4(4k^2+1) = 0.$$

The expression $p^2q - 4q^2 = a^4k^6(4k^2+1)^2$ is now automatically a rational square.

THEOREM 2. *The equation*[2]

$$x^4 + ak(4k^2+1)x^2 + a^2k^4(4k^2+1) = 0, \tag{16}$$

where a and k take on all rational values, is cyclic if it is irreducible, and all cyclic quartic fields are so obtained.

The automorphisms (12) become

$$\alpha, \quad \alpha' = \theta(\alpha) = \frac{\alpha^3 + ak(2k^2+1)\alpha}{ak^2}, \quad \alpha'' = -\alpha, \quad \alpha''' = -\theta(\alpha). \tag{17}$$

Garver [3] developed the cyclic quartic equation

$$f(x) = x^4 - 2g(1+e^2)x^2 + e^2g^2(1+e^2) = 0.$$

If we substitute $-2g = ak$ and $e = 2k$, this is identical with (16).

Albert [1] states that every cyclic quartic field $F(x)$ over F is generated by a quantity x satisfying

$$x^2 = v(u - \tau), \qquad u^2 = \tau = 1 + e^2$$

with the generating automorphism $x' = (1+u)x/e$. This representation of the cyclic quartic will also be proved equivalent to (16).

If F is the rational field, and if α is a root of an irreducible equation with coefficients in the algebraic subfield $F(\beta)$ of $F(\alpha)$, Williamson [10] has given a practicable method for finding an equation with coefficients in F satisfied by α. Let α satisfy the irreducible equation $a(x) = 0$ of degree m with coefficients in $F(\beta)$ and let β satisfy the irreducible equation $b(x) = 0$ of degree n with coefficients in F. Let A be the companion matrix of $a(x) = 0$ with elements in $F(\beta)$, and let B be the companion matrix of $b(x) = 0$ with rational coefficients. Let C be the $mn \times mn$ matrix obtained by replacing β by B in every element of A. Then $|C - \lambda I| = 0$ has rational coefficients and is satisfied by α.

This theorem is directly applicable to show that Albert's representation is equivalent to (16). The characteristic equation of B, $u^2 = \tau$, defines a quadratic subfield $F(u)$ of the desired quartic field $F(\alpha)$; $f(x) = 0$, which defines

[2] See [3].

$F(\alpha)$, is quadratic over this subfield and is expressed as $x^2 = v(u - \tau)$, a polynomial with coefficients in $F(u)$. When in the matrix

$$A = \begin{Vmatrix} 0 & 1 \\ v(u-\tau) & 0 \end{Vmatrix}$$

u is replaced by its matric equivalent

$$B = \begin{Vmatrix} 0 & 1 \\ \tau & 0 \end{Vmatrix},$$

the new 4×4 matrix has for its characteristic equation $f(x) = 0$, a quartic equation with coefficients in F which is cyclic over $F(u)$, and, since $F(u)$ is cyclic over F, the roots of $f(x) = 0$ will form a cyclic group of order 4. Thus

$$C = \begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -v\tau & v & 0 & 0 \\ v\tau & -v\tau & 0 & 0 \end{Vmatrix},$$

and $| C - xI | = f(x) = x^4 + 2v(1 + e^2)x^2 + v^2e^2(1 + e^2) = 0$, an equation which is equivalent to (16).

5. Normal quartics with the 4-group. From (15),

$$a = \pm \frac{1}{q^{\frac12}}, \qquad b = \pm \frac{p}{q^{\frac12}}, \qquad p = \frac{b}{a}, \qquad q = \frac{1}{a^2}.$$

Then (11) becomes

$$f(x) = x^4 + \frac{b}{a}x^2 + \frac{1}{a^2} = 0, \qquad a \neq 0,$$

which can be transformed by $x = ay$ into

$$y^4 + aby^2 + a^2 = 0.$$

THEOREM 3. *The equation*[3]

$$x^4 + abx^2 + a^2 = 0, \tag{18}$$

where a and b take on all rational values, is normal with the 4-group as its automorphism group provided it is irreducible; all normal quartic fields with the 4-group are thus obtained.

The automorphisms (12) are

$$\alpha, \qquad \alpha' = \theta(\alpha) = \frac{\alpha^3 + ab\alpha}{a}, \qquad \alpha'' = -\alpha, \qquad \alpha''' = -\theta(\alpha). \tag{19}$$

[3] See [3].

A second derivation of (18) will be given by a method which can be extended to equations of higher degree. Since the 4-group is the direct product of two quadratic subgroups, the normal field with the 4-group is the direct product of two quadratic subfields.

Let $f_1(x) = x^2 + ax + b = 0$, with a and b in F, be irreducible in F and let A be a matrix having $f_1(x) = 0$ as its characteristic equation. Let $f_2(y) = y^2 + c$, with c rational, be irreducible in F and let B be a matrix having $f_2(y) = 0$ as its characteristic equation. Then we may take

$$A = \begin{Vmatrix} 0 & 1 \\ -b & -a \end{Vmatrix}, \qquad B = \begin{Vmatrix} 0 & 1 \\ -c & 0 \end{Vmatrix}.$$

Now form the direct product of the matrices $A \times B = C$, that is,

$$C = \begin{Vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -b & -a \\ 0 & -c & 0 & 0 \\ cb & ca & 0 & 0 \end{Vmatrix}.$$

The characteristic equation of C is

$$(20) \qquad | C - \lambda I | = \lambda^4 + c(a^2 - 2b)\lambda^2 + b^2c^2 = 0,$$

which is another representation of all quartics with the 4-group.

This in fact gives the same totality of equations as $x^4 + efx^2 + f^2 = 0$ for we may let $f = bc$, $e = (a^2 - 2b)/b$, $a = f$, $b = f^2/(2 + e)$, $c = (2 + e)/f$. No denominator is zero if (20) is irreducible.

6. **Cyclic sextics.**[4] Consider a normal equation of degree 6 which is either cyclic or has the symmetric group of order 6 as its Galois group. In either case the group is imprimitive and the sextic can be put [11] into the form

$$(21) \qquad f(x) = x^6 + px^4 + qx^2 + r = 0$$

with p, q, and r rational.

Since the cyclic group of order 6 is the direct product of two cyclic subgroups of orders two and three, respectively, the cyclic field of degree 6 is the direct product of two cyclic subfields of degrees two and three. From Theorem 1, (9) defines a cyclic cubic field if $(2b + 1)^2 = -2e - 3$. Let A be a matrix which has as its characteristic equation $f_1(x) = x^3 + 6ex - 4e(2b + 1) = 0$, e.g.,

$$A = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4e(2b+1) & -6e & 0 \end{Vmatrix}.$$

[4] See [2].

Let $f_2(y) = y^2 - d = 0$, with d rational but not a square, define a quadratic field and let B be a matrix which has $f_2(y) = 0$ as its characteristic equation, e.g.,

$$B = \begin{Vmatrix} 0 & 1 \\ d & 0 \end{Vmatrix}.$$

The direct product $A \times B$ of the matrices is C, where

$$C = \begin{Vmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & d & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & d & 0 \\ 0 & 4e(2b+1) & 0 & -6e & 0 & 0 \\ 4ed(2b+1) & 0 & -6ed & 0 & 0 & 0 \end{Vmatrix}.$$

The characteristic equation of C is given below as (22).

THEOREM 4. *The equation*

$$f(\lambda) = \lambda^6 + 12ed\lambda^4 + 36e^2d^2\lambda^2 + 16e^2d^3(2e+3) = 0, \tag{22}$$

where d is any rational number not a square and $(2b+1)^2 = -2e - 3$, is cyclic if it is irreducible, and all cyclic sextic fields are so obtained.

The automorphism group of (22) is the direct product of the automorphism groups of its subfields. From (10) and the automorphisms β and $\beta' = -\beta$ of $f_2(y) = 0$ we form the 6 products

$$\alpha\beta, \qquad \alpha'\beta' = -\tfrac{1}{2}\alpha^2\beta - b\alpha\beta - 2e\beta, \qquad \alpha''\beta' = \tfrac{1}{2}\alpha^2\beta + (b+1)\alpha\beta + 2e\beta,$$

$$\alpha\beta' = -\alpha\beta, \qquad \alpha'\beta = -\alpha'\beta', \qquad \alpha''\beta = -\alpha''\beta'.$$

Let $\gamma = \alpha\beta$ be one root of (22). Since $\gamma^2 = d\alpha^2$ and $\gamma^3 = \alpha^3\beta^3 = -6ed\alpha\beta + 4ed\beta$, we have

$$\beta = \frac{\gamma^3 + 6ed\gamma}{4ed(2b+1)}, \qquad \alpha^2\beta = \frac{\gamma^5 + 6ed\gamma^3}{4ed^2(2b+1)}.$$

The denominators of β and $\alpha^2\beta$ can not be zero if (22) is normal. Hence the automorphisms of (22) are

$$\begin{gathered} \gamma, \qquad \gamma' = \theta_1(\gamma) = -\frac{\gamma^5 + 10ed\gamma^3 + 8ed^2(2e - b - 2)\gamma}{8ed^2(2b+1)}, \\ \gamma'' = \theta_2(\gamma) = \frac{\gamma^5 + 10ed\gamma^3 + 8ed^2(2e + b - 1)\gamma}{8ed^2(2b+1)}, \\ \gamma''' = -\gamma, \qquad \gamma^{\text{iv}} = -\theta_1(\gamma), \qquad \gamma^{\text{v}} = -\theta_2(\gamma). \end{gathered} \tag{23}$$

The set (23) forms the cyclic group of order 6.

7. **Normal sextics with the symmetric group.** The cubic equation $f_1(x) = x^3 + px + q = 0$ is not normal in the rational field F when the discriminant d

is not a square, but $f_2(y) = y^2 - d = 0$ defines a quadratic field in which $f_1(x) = 0$ is normal. The 6 products of the roots (8) and the roots β and $\beta' = -\beta$ of $f_2(y) = 0$ are the rational functions of α and β

$$\begin{aligned} &\alpha\beta, \qquad \alpha'\beta = \tfrac{1}{2}(6p\alpha^2 - 9q\alpha - \alpha\beta + 4p^2), \\ (24) \qquad &\alpha''\beta = \tfrac{1}{2}(-6p\alpha^2 + 9q\alpha - \alpha\beta - 4p^2), \qquad \alpha\beta' = -\alpha\beta, \\ &\alpha'\beta' = -\alpha'\beta, \qquad \alpha''\beta' = -\alpha''\beta, \end{aligned}$$

and are the 6 roots of an equation $f(x, y) = 0$, where the x and y are determined by $f_1(x) = 0$ and $f_2(y) = 0$, respectively. If we let A be a matrix which has the characteristic equation $f_1(x) = 0$ and B a matrix with the characteristic equation $f_2(y) = 0$, by a theorem of Stéphanos [6] the characteristic roots of the direct product $A \times B = C$ are the products $\alpha^{(i)}\beta^{(j)}$ of a characteristic root of A and one of B. If the direct product is formed, the characteristic equation of C becomes

$$(25) \qquad | C - \lambda I | = f(\lambda) = \lambda^6 + 2pd\lambda^4 + p^2d^2\lambda^2 - q^2d^3 = 0.$$

Let $\gamma = \alpha\beta$ be one root of (25). Then since $\gamma^2 = d\alpha^2$ and $\gamma^4 = -d^2p\alpha^2 - qd^2\alpha$, $\alpha^2 = \gamma^2/d$, $\alpha = -(\gamma^4 + pd\gamma^2)/qd^2$. Neither q nor d can be zero if (25) is to be irreducible. The conjugates of γ become from (24)

$$\begin{aligned} &\gamma, \qquad \gamma' = \alpha'\beta = \theta_1(\gamma) = \frac{9\gamma^4 + 15pd\gamma^2 - d^2\gamma + 4p^2d^2}{2d^2}, \\ (26) \qquad &\gamma'' = \alpha''\beta = \theta_2(\gamma) = -\frac{9\gamma^4 + 15pd\gamma^2 + d^2\gamma + 4p^2d^2}{2d^2}, \\ &\gamma''' = -\alpha\beta = -\gamma, \qquad \gamma^{\text{iv}} = -\theta_1(\gamma), \qquad \gamma^{\text{v}} = -\theta_2(\gamma). \end{aligned}$$

Hence (25) is a normal equation in F. The set (26) represents the Galois group of (25). It can be verified that (26) represents the symmetric group of order 6.

THEOREM 5. *The equation*

$$(25) \qquad f(\lambda) = \lambda^6 + 2pd\lambda^4 + p^2d^2\lambda^2 - q^2d^3 = 0,$$

where p and q take on all rational values such that $d = -4p^3 - 27q^2$ is not a square, is normal with the symmetric group of order 6 *as its automorphism group provided it is irreducible; all normal sextic fields with the symmetric group are so obtained.*

8. **Normal octics with the group** $G(2, 2, 2)$. A normal Abelian equation of degree 8 has either the group $G(2, 2, 2)$ with no cyclic subgroup of order greater than 2, the group $G(4, 2)$ with a cyclic subgroup of order 4, or is cyclic. Each of these groups is imprimitive and the equation can be put [11] into the form

$$(27) \qquad f(x) = x^8 + px^6 + qx^4 + rx^2 + s = 0,$$

where p, q, r and s are rational.

Since the group $G(2, 2, 2)$ is the direct product of the 4-group and a group of order 2, the normal field with this group as its automorphism group is the direct product of a normal quartic field with the 4-group and a quadratic field.

Let $f_1(x) = x^4 + abx^2 + a^2 = 0$, an equation which from Theorem 3 defines a normal quartic field with the 4-group if it is irreducible, and let A be a matrix which has $f_1(x) = 0$ as its characteristic equation. Let $f_2(y) = y^2 + dy + e = 0$ be irreducible in F, thus defining a quadratic field, and let B be a matrix which has $f_2(y) = 0$ as its characteristic equation. Form the direct product $A \times B = C$ as in the preceding paragraphs. Then the characteristic equation of C is given below as (28).

Theorem 6. *The equation*

$$f(\lambda) = \lambda^8 + ab\rho\lambda^6 + a^2(b^2e^2 + \rho^2 - 2e^2)\lambda^4 + a^3be^2\rho\lambda^2 + a^4e^4 = 0, \tag{28}$$

where $\rho = d^2 - 2e$ *and* a, b, d, e *range over all rational numbers, is normal with the Galois group* $G(2, 2, 2)$ *provided it is irreducible; all normal octic fields with the group* $G(2, 2, 2)$ *are thus obtained.*

The Galois group of (28) is the direct product of (19) and the Galois group of $f_2(y) = 0$ consisting of β and $\beta' = -d - \beta$. Since (28) is a normal equation, every element of its Galois group is a rational polynomial in one root γ. By steps entirely similar to those used in §7 we find the automorphisms of (28) to be

$$\begin{aligned}
&\gamma, \quad \gamma' = \theta_1(\gamma) = \frac{\rho\gamma^7 + ab(\rho^2 - e^2)\gamma^5 + a^2\rho(\rho^2 - e^2)\gamma^3 + a^3be^2(\rho^2 - b^2e^2 + e^2)\gamma}{a^3e^2(\rho^2 - b^2e^2)},\\
&\gamma'' = \theta_2(\gamma) = \frac{\gamma^7 + ab\rho\gamma^5 + a^2(b^2e^2 + \rho^2 - 2e^2)\gamma^3 + a^3be^2\rho\gamma}{a^3e^3},\\
&\gamma''' = \theta_3(\gamma) = \frac{b\gamma^7 + a\rho(b^2 - 1)\gamma^5 + a^2be^2(b^2 - 1)\gamma^3 + a^3\rho(b^2e^2 - \rho^2 + e^2)\gamma}{a^3e(b^2e^2 - \rho^2)},\\
&\gamma^{\text{iv}} = -\gamma, \qquad \gamma^{\text{v}} = -\theta_1(\gamma), \qquad \gamma^{\text{vi}} = -\theta_2(\gamma), \qquad \gamma^{\text{vii}} = -\theta_3(\gamma).
\end{aligned} \tag{29}$$

No denominator can be zero if (28) is normal and $\rho = d^2 - 2e$. The set (29) forms the group $G(2, 2, 2)$.

9. **Normal octics with the group** $G(4, 2)$. Since the group $G(4, 2)$ is the direct product of the cyclic group of order 4 and a group of order 2, we can construct the normal field with the automorphism group $G(4, 2)$ by steps similar to those used in the preceding case.

Let $f_1(x) = x^4 + ak(4k^2 + 1)x^2 + a^2k^4(4k^2 + 1) = 0$, an equation which by Theorem 2 is cyclic if it is irreducible, and let A be a matrix which has $f_1(x) = 0$ as its characteristic equation. Then let $f_2(y) = y^2 + dy + e = 0$ be irreducible in F and let B be a matrix having $f_2(y) = 0$ as its characteristic equation. Again, the direct product $A \times B = C$ is formed and the characteristic equation of C is given below as (30).

THEOREM 7. *The equation*

$$(30)\qquad f(\lambda) = \lambda^8 + ak\tau\rho\lambda^6 + a^2k^2\tau[\tau e^2 + k^2(\rho^2 - 2e^2)]\lambda^4 + a^3k^5e^2\rho\tau^2\lambda^2 + a^4e^4k^8\tau^2 = 0,$$

where $\rho = d^2 - 2e$, $\tau = 4k^2 + 1$ and a, d, e, k take on all rational values such that neither $d^2 - 4e$ nor $4k^2 + 1$ is a square, is normal with the Galois group $G(4, 2)$ provided it is irreducible; all normal octic fields with the group $G(4, 2)$ are so obtained.

The Galois group of (30) is the direct product of (17) and the automorphisms of $f_2(y) = 0$ and can be represented as functions of one root of (30). The automorphisms are

$$(31)\qquad \begin{array}{cccc} \gamma, & \gamma' = \theta_1(\gamma), & \gamma'' = \theta_2(\gamma), & \gamma''' = \theta_3(\gamma), \\ \gamma^{iv} = -\gamma, & \gamma^{v} = -\theta_1(\gamma), & \gamma^{vi} = -\theta_2(\gamma), & \gamma^{vii} = -\theta_3(\gamma), \end{array}$$

where

$$\begin{aligned} -ke^{-2}\phi\theta_1 &= \gamma^7 + ak(3k^2+1)\rho\gamma^5 + a^2e^2k^2\tau(3k^2+1)\gamma^3 \\ &\qquad + a^3k^5\rho\tau[\tau e^2 - k^2(\rho^2 - e^2)]\gamma, \\ e^{-1}\phi\theta_2 &= \rho\gamma^7 + ak\tau(\rho^2 - e^2)\gamma^5 + a^2k^4\tau\rho(\rho^2 - e^2)\gamma^3 \\ &\qquad + a^3k^3e^2\tau[k^2\rho^2(2k^2+1) - \tau e^2(k^2+1)]\gamma, \\ k^2\phi\theta_3 &= [\rho^2k^2 - e^2(2k^2+1)]\gamma^7 + ak\rho[k^2\tau(\rho^2 - e^2) - e^2(3k^2+1)(2k^2+1)]\gamma^5 \\ &\qquad + \tau a^2k^2[e^4(\tau + 2k^4) - k^4\rho^2(\rho^2 - 2e^2)]\gamma^3 - a^3\rho k^5e^2\tau[\tau e^2(k^2+1) \\ &\qquad - k^2(2k^2+1)(\rho^2 - e^2)]\gamma, \end{aligned}$$

$$\rho = d^2 - 2e, \qquad \tau = 4k^2 + 1, \qquad \phi = a^3k^4e^3\tau(k^2\rho^2 - \tau e^2).$$

The coefficients of the θ's can not be zero if (30) is normal. The set (31) forms the group $G(4, 2)$.

10. **Cyclic octics.** In his paper on cyclic fields of degree 8, Albert [1] states the theorem:

Every cyclic field $F(x)$ of degree 8 over F is generated by a number x satisfying

$$x^2 = a, \qquad x' = \beta x,$$

with

$$a = \frac{\beta_1' u}{ve\tau}[veu\delta_1 - v(u - \tau)\delta_1' + (\delta_1\delta_1' - veu)y_0],$$

and

$$\beta = \frac{\beta_1}{ve\tau}\left[ve\tau - \frac{\delta_1}{e}(u + \tau)y_0\right],$$

where $y_0^2 = v(u - \tau)$, $u^2 = \tau = 1 + e^2$ in F, $\delta_1 = \xi_1 + \xi_2 u$, $\beta_1 = \xi_3 + \xi_4 u$, with $\xi_1, \xi_2, \xi_3, \xi_4$ in F such that $v \neq 0$ in F and $\delta_1 \neq 0$, $\beta_1 \neq 0$; and if

$$\lambda = \xi_3^2 - \xi_4^2\tau = \beta_1\beta_1'$$

in F, then

$$\delta_1^2 = v(u - \tau + \lambda^{-1}eu).$$

Since here we are interested in cyclic fields of degree 8 over the rational field F, it is necessary to find rational solutions for δ_1 and β_1 and construct an equation which fulfills the conditions of the theorem.

By definition $\delta_1 = \xi_1 + \xi_2 u$, so that $\xi_1^2 + 2\xi_1\xi_2 u + \tau\xi_2^2 = v(u - \tau + \lambda^{-1}eu)$. We must find rational solutions for the three relations

$$-v\tau = \xi_1 + \tau\xi_2^2, \qquad v(1 + \lambda^{-1}e) = 2\xi_1\xi_2, \qquad \lambda = \xi_3^2 - \xi_4^2\tau.$$

Since e is rational, $\tau > 0$ and hence $v < 0$. We can assume that v has no square factor except 1. If, in particular, $v = -1$, then $\xi_1 = 0$ and $\xi_2 = 1$ satisfy the first relation. The second relation becomes $1 + \lambda^{-1}e = 0$, so that $\lambda = -e$ and is rational. If we let $e = b^2$, the third relation has the solutions $\xi_3 = -\xi_4 = 1/b$. Thus $\beta_1 = (1 - u)/b$ and $\delta_1 = u$, and to every rational b there will correspond a cyclic field of degree 8. This field will be generated by an x satisfying

$$x^2 = a = \tau b^{-1} + (1 - b^2)ub^{-1} + [\tau - b^2 + (1 - b^2)u]y_0 b^{-3}.$$

Let A be a matrix which has $x^2 = a$ for its characteristic equation, e.g.,

$$A = \begin{Vmatrix} 0 & 1 \\ a & 0 \end{Vmatrix}.$$

Since $y_0^2 = \tau - u$ is the characteristic equation of the matrix B, where

$$B = \begin{Vmatrix} 0 & 1 \\ \tau - u & 0 \end{Vmatrix},$$

by Williamson's theorem [10] we can replace y_0 in A by B and obtain

$$A = \begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \dfrac{\tau + (1 - b^2)u}{b} & \dfrac{\tau - b^2 + (1 - b^2)u}{b^3} & 0 & 0 \\ \tau b - b^3 u & \dfrac{\tau + (1 - b^2)u}{b} & 0 & 0 \end{Vmatrix}.$$

But $\tau = 1 + b^4 = u^2$ is the characteristic equation of the matrix

$$C = \begin{Vmatrix} 0 & 1 \\ \tau & 0 \end{Vmatrix}.$$

If in A we replace u by the matrix C, we have A expressed as an 8×8 matrix with the characteristic equation given below as (32).

THEOREM 8. *The equation*

$$f(x) = b^4x^8 - 4b\tau x^6 + (b^6 + 8b^2 + 1)\tau x^4 - (b^6 + 3b^2 + 4)b\tau x^4 + b^2\tau = 0, \tag{32}$$

where $\tau = 1 + b^4$ *and* b *is rational, is cyclic if it is irreducible.*

When $\beta_1 = (1 - u)/b$ and $\delta_1 = u$, $\beta = (1 - u - y_0)/b$ and the automorphisms of (32) are

$$\alpha, \quad \alpha' = \beta\alpha, \quad \alpha'' = \beta'\beta\alpha = [-b^2 + u + (b^2 - 1)y_0 b^{-2} - uy_0 b^{-2}]\alpha,$$
$$\alpha''' = \beta''\beta'\beta\alpha = [-b + y_0 b^{-1}]\alpha, \quad \alpha^{\text{iv}} = \beta'''\beta''\beta'\beta\alpha = -\alpha,$$
$$\alpha^{\text{v}} = -\alpha', \quad \alpha^{\text{vi}} = -\alpha'', \quad \alpha^{\text{vii}} = -\alpha'''.$$

The automorphisms form a cyclic group of order 8. They can be expressed explicitly as polynomials in α with rational coefficients. These polynomials are

$$\begin{aligned}
\alpha' &= a_{11}\alpha^7 + a_{21}\alpha^5 + a_{31}\alpha^3 + a_{41}\alpha,\\
\alpha'' &= a_{12}\alpha^7 + a_{22}\alpha^5 + a_{32}\alpha^3 + a_{42}\alpha,\\
\alpha''' &= a_{13}\alpha^7 + a_{23}\alpha^5 + a_{33}\alpha^3 + a_{43}\alpha,\\
\alpha^{\text{iv}} &= -\alpha, \quad \alpha^{\text{v}} = -\alpha', \quad \alpha^{\text{vi}} = -\alpha'', \quad \alpha^{\text{vii}} = -\alpha''',
\end{aligned} \tag{33}$$

where

$$\Delta a_{11} = -b^2(1 - b^2 - 3b^4 + 4b^6 - b^8),$$
$$\Delta a_{21} = b^3(6 - 41b^2 + 68b^4 - 69b^6 + 63b^8 - 26b^{10}),$$
$$\Delta a_{31} = 14 - 54b^2 + 179b^4 + 312b^6 + 430b^8 - 482b^{10} + 359b^{12} - 228b^{14} + 89b^{16},$$
$$b\Delta a_{41} = -8 + 25b^2 - 84b^4 + 161b^6 - 282b^8 + 347b^{10} - 303b^{12} + 256b^{14} - 125b^{16} + 45b^{18} - 26b^{20},$$
$$\Delta a_{12} = b^3(1 - 5b^4 + 5b^6),$$
$$\Delta a_{22} = -b^4(2 - 42b^2 + 65b^4 - 54b^6 + 53b^8 - 9b^{10} - 8b^{12}),$$
$$\Delta a_{32} = -b(15 - 69b^2 + 250b^4 - 450b^6 + 647b^8 - 745b^{10} + 557b^{12} - 363b^{14} + 144b^{16}),$$
$$\Delta a_{42} = 6 - 87b^2 + 283b^4 - 518b^6 + 805b^8 - 953b^{10} + 887b^{12} - 773b^{14} + 500b^{16} - 274b^{18} + 141b^{20} - 24b^{22},$$
$$\Delta a_{13} = b^2(1 - b^2 - b^4 + 2b^6 - b^8),$$
$$\Delta a_{23} = b^3(-5 + 15b^2 - 23b^4 + 24b^6 - 19b^8 + 8b^{10}),$$
$$\Delta a_{33} = -14 + 50b^2 - 101b^4 + 144b^6 - 158b^8 + 140b^{10} - 89b^{12} + 46b^{14} - 17b^{16},$$
$$b\Delta a_{43} = 8 - 24b^2 + 74b^4 - 129b^6 + 146b^8 - 124b^{10} + 66b^{12} - 25b^{14} - 2b^{16} - b^{18} + 8b^{20},$$

with

$$\Delta = 1 - 26b^2 + 54b^4 - 48b^6 + 19b^8 + 24b^{10} - 51b^{12} + 46b^{14} - 18b^{16}.$$

Bibliography

1. A. A. Albert, *Cyclic fields of degree* 8, Transactions of the American Mathematical Society, vol. 35(1933), p. 949.

2. S. Breuer, *Zyklische Gleichungen* 6 *Grades und Minimal Basis*, Mathematische Annalen, vol. 86(1922), pp. 108–113.

3. R. Garver, *Quartic equations with certain groups*, Annals of Mathematics, (2), vol. 30(1929), pp. 47–51.

4. G. Bucht, *Über einige algebraische Körper octen Grades*, Arkiv för Matematik, Astronomi och Fysik, vol. 6(1910), No. 30.

5. R. Hull, *A determination of all cyclotomic quintic fields*, Annals of Mathematics, (2), vol. 36(1935), pp. 366–372.

6. C. C. MacDuffee, *Theory of Matrices*, Berlin, 1933, p. 84.

7. E. Noether, *Gleichungen mit vorgeschriebener Gruppe*, Mathematische Annalen, vol. 78(1917), pp. 221–229.

8. F. Seidelmann, *Die Gesamtheit der kubischen und biquadratischen Gleichungen mit Affekt*, Mathematische Annalen, vol. 78(1917), pp. 230–233.

9. J. A. Serret, *Cours d'Algèbre Supérieure*, vol. 2, p. 467.

10. J. Williamson, *The latent roots of a matrix of special type*, Bulletin of the American Mathematical Society, vol. 37(1931), p. 585.

11. A. Winan, *Über gewisse imprimitive Gleichungen*, Arkiv för Matematik, Astronomi och Fysik, vol. 3(1906), No. 28.

University of Wisconsin.

AN AFFINE INVARIANT OF CONVEX REGIONS

By Olin B. Ader

1. **Introduction.** Let R denote a closed and bounded three-dimensional convex region. Let D be the greatest diameter (Durchmesser), and let Δ be the smallest diameter (Dicke) of R.[1] We speak of convex regions as belonging to different *classes*, a single class consisting of those regions which can be derived from one another by affine transformations.

The problem treated in this paper is that of finding the maximum for all classes C of the minimum values of D/Δ for all regions l belonging to C; i.e., to find $\max_{C} \min_{l \subset C} D/\Delta$. This problem has been solved originally by F. Behrend[2] and subsequently by F. John[3] for the two-dimensional case. They find $\max_{C} \min_{l \subset C} D/\Delta = \sqrt{2}$, and the class of parallelograms is the only class which attains this maximum. In the three-dimensional case, we find that the constant is $\sqrt{3}$, and that, of convex regions having a center, there are *two* classes which attain this maximum, namely, the parallelepipeds and the octahedrons.[4] If we measure *similarity of a region to a sphere* by the value D/Δ, our result states how similar to a sphere a convex region can be made by an affine transformation.

2. **A necessary condition for minimal regions.** Let us call a region l_1 of the class C for which D/Δ is a minimum a *minimal region* of the class. We first derive a necessary condition that the convex region l_1 shall be a minimal region.

We suppose that the convex region R has a *center*, i.e., R is symmetrical with respect to some point. We will consider later convex regions without center, a modification which presents no difficulty. In the present case, $D/\Delta = 1$ only for spheres, and we note that $\min_{C} \min_{l \subset C} D/\Delta = 1$ is attained only by the class of ellipsoids. This is a trivial result.

Take the center of R as the origin of a rectangular system of coördinates. The distance from the center to any point on the boundary of R we shall call a *radius*, r, of R. Then, $\frac{1}{2}D = \max r$, and $\frac{1}{2}\Delta = \min r$. We apply to the

Received October 15, 1937.

[1] A diameter of R is the distance apart of two parallel planes of support. For definitions, see Bonnesen-Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Mathematik, vol. 3, No. 1, Berlin, 1934, p. 37 ff.

[2] F. Behrend, *Über einige Affininvarianten konvexer Bereiche*, Math. Annalen, vol. 113, pp. 713–747.

[3] F. John, *Moments of inertia of convex regions*, this Journal, vol. 2(1936), pp. 447–452.

[4] A proof that the constant in question is $\leqq \frac{1}{3}\sqrt{30}$ is given by John, loc. cit. The methods used in the present paper are more closely related to those of Behrend.

boundary points x_i of R the affine transformation with non-vanishing determinant

$$x'_i = x_i + \sum_j \epsilon\, a_{ij} x_j \qquad (i, j = 1, 2, 3),$$

where ϵ assumes sufficiently small positive values. Any radius, $r(x)$, of the region then becomes a radius, $r(x')$, of the transformed region

$$r^2(x') = \sum x_i'^2 = \sum (x_i^2 + 2\,\epsilon\, a_{ij} x_i x_j) + O(\epsilon^2) = r^2(x)\left[1 + \frac{2\,\epsilon\, Q(x)}{r^2(x)}\right] + O(\epsilon^2),$$

where $O(\epsilon^2)$ stands for a quantity whose value remains smaller than a fixed constant multiplied by ϵ^2. $Q(x)$ stands for the quadratic form $\sum_{i,j} a_{ij} x_i x_j$.

Let us assume that R is a minimal region, so that the value of D/Δ is not reduced by any affine transformation. Let $r(y)$ be the length of a radius of R which is transformed into a greatest radius of R' of length $\frac{1}{2}D'$, and let $r(z)$ be the length of a radius which is transformed into the length $\frac{1}{2}\Delta'$. Then y_i, z_i $(i = 1, 2, 3)$ are the coördinates of the end points of the radii of R of lengths $r(y)$ and $r(z)$, respectively. We now have

$$\text{(1)} \qquad \frac{D'^2}{\Delta'^2} = \frac{r^2(y)[1 + 2\epsilon Q(y)/r^2(y)] + O(\epsilon^2)}{r^2(z)[1 + 2\epsilon Q(z)/r^2(z)] + O(\epsilon^2)} \geqq \frac{D^2}{\Delta^2}.$$

We know that $2r(y) \leqq D$ and $2r(z) \geqq \Delta$, since D and Δ are, respectively, the greatest and smallest diameters of R. Hence,

$$\frac{\frac{1}{4}D^2[1 + 2\epsilon Q(y)/r^2(y)] + O(\epsilon^2)}{\frac{1}{4}\Delta^2[1 + 2\epsilon Q(z)/r^2(z)] + O(\epsilon^2)} \geqq \frac{D^2}{\Delta^2},$$

and

$$\frac{[1 + 2\epsilon Q(y)/r^2(y)] + O(\epsilon^2)}{[1 + 2\epsilon Q(z)/r^2(z)] + O(\epsilon^2)} \geqq 1,$$

or

$$\frac{\epsilon Q(y)}{r^2(y)} \geqq \frac{\epsilon Q(z)}{r^2(z)} + O(\epsilon^2).$$

Since $O(\epsilon^2) = \Theta\epsilon^2$, where $|\,\Theta\,| \leqq c$, a constant, we have

$$\text{(2)} \qquad \frac{Q(y)}{r^2(y)} \geqq \frac{Q(z)}{r^2(z)} + \Theta\epsilon.$$

The coefficients a_{ij} in $Q(y)$ and $Q(z)$ are fixed and independent of ϵ. D' and Δ' depend on ϵ, and so must their originals of length $2r(y)$ and $2r(z)$, which have the coördinates y_i and z_i. Let $\bar{y}_i$ and $\bar{z}_i$ be the coördinates of the end points of the greatest and smallest diameters, respectively, in the region R. y_i and z_i depend on ϵ. A suitable subsequence of these will converge toward $\bar{y}_i$

and $\bar{z}_i$, respectively, for ϵ approaching zero. By (1), we have in the limit as ϵ approaches zero

$$\frac{r^2(\bar{y})}{r^2(\bar{z})} \geqq \frac{D^2}{\Delta^2}.$$

Since $r(\bar{y})$ and $r(\bar{z})$ are radii of R, only the equality sign may hold, and $2r(\bar{y}) = D$, and $2r(\bar{z}) = \Delta$. Equation (2) becomes in the limit

$$\frac{Q(\bar{y})}{r^2(\bar{y})} \geqq \frac{Q(\bar{z})}{r^2(\bar{z})}$$

for every quadratic form Q. The coördinates $\bar{y}_i$, $\bar{z}_i$ of the end points of the radii may be written, respectively, as $r(\bar{y}) \cos \alpha_i$ and $r(\bar{z}) \cos \beta_i$. Hence, we say that at least for some D and some Δ

$$Q(\cos \alpha) \geqq Q(\cos \beta),$$

where the α's are the angles associated with D and the β's are the angles associated with Δ. To simplify the notation, let $\cos \alpha \equiv y$ and $\cos \beta \equiv z$. Given a quadratic form Q and a constant c, there cannot be a distribution of the D's and Δ's in space such that

$Q(y) \leqq c$ for all greatest diameters,

$Q(z) > c$ for all smallest diameters.

The equality sign is taken in the first equation, although it may as well be taken in the second, but may not be taken in both equations. These conditions are equivalent to the conditions

$$Q_1(y) < 0, \quad Q_1(z) > 0,$$

where the coefficients a_{ij} in the new quadratic are still regarded as arbitrary. The locus of $Q_1 = 0$ is a quadratic cone in space. This leads us to the following theorem which we shall refer to as the *condition of non-separation*:

THEOREM 1. *In a convex region which is minimal with respect to D/Δ under affine transformations, there can be no quadratic cone with vertex at the center of the region and having all the greatest (smallest) diameters interior to the region it bounds or on its surface and having all the smallest (greatest) diameters exterior to this region.*

3. **Number of D's and Δ's.** Let y^s $(s = 1, \cdots, s')$ be, for each s, coördinates associated with different D's, s ranging over their number, and let z^t $(t = 1, \cdots, t')$ be, for each t, coördinates associated with different Δ's, t ranging over their number. The coördinates y^s and z^t can now be regarded as fixed by these diameters, which we know to be not separated by any quadratic cone. We say, then, that the inequalities

$$(3) \qquad \begin{aligned} -Q_1(y) &\equiv -\sum_{i,j} a_{ij} y_i^s y_j^s > 0 \quad \text{for all } s, \\ Q_1(z) &\equiv \sum_{i,j} a_{ij} z_i^t z_j^t > 0 \quad \text{for all } t \end{aligned}$$

must be incompatible for all different fixed y^s and z^t, which are arbitrary except for the condition of non-separation, and for every quadratic form Q. That is, these inequalities must have no solution in the six quantities a_{ij} regarded as the variables. We now make use of the theorem:[5]

The system $\sum_{k=1}^{m} u_{ik}a_k > 0$ $(i = 1, \cdots, n)$ *has a solution if and only if the origin is exterior to the convex figure associated with the representative points, i.e., the n points* $(u_{i1}, \cdots, u_{im})$ *located in m-space.*

(We note that n may be infinite.) In order that the system (3) have no solution, then, we require that the origin be in the interior or on the boundary of the convex figure associated with the representative points, which are $-y_i^s y_j^s$ and $z_i^t z_j^t$, or $s' + t'$ points in a six-space. It is then possible to choose a subset of the representative points consisting of at most 7 points, such that the origin lies in the interior or on the boundary of the simplex with these points as vertices.[6] Thus, one may choose a subsystem of no more than 7 of the equations (3) which are incompatible; the corresponding greatest or smallest diameters then form a system of no more than 7 diameters already satisfying the non-separation condition. *In what follows we restrict ourselves to the consideration of this system of diameters.* We then have $s' + t' \leqq 7$.[7]

4. **Considerations in a projective plane.** If we regard the traces of the end points of the s' D's and the t' Δ's in a projective plane, we may say that no conic separates these traces. It follows from $s' + t' \leqq 7$, that either $s' \leqq 3$ or $t' \leqq 3$; we shall restrict ourselves to $s' \leqq 3$, i.e., the case of no more than 3 D's, since the dual case, $t' \leqq 3$, can be easily reduced to the other one as shown in §5. There must be more than one greatest diameter, for this one would be separable from the smallest diameters by some sufficiently small circle about it as a center. If there are just two D's, we have three cases to consider. (a) If no Δ lies on the line between the two D's, then a sufficiently narrow ellipse having the two D's at the end points of its major axis will separate the D's and Δ's. (b) If no Δ lies on the line extending outward from the two D's, then a sufficiently narrow hyperbola, having the two D's as its vertices and this line as its axis, will separate the D's and Δ's. (c) If a Δ lies on each of the line segments mentioned in (a) and (b), we then have in the three-dimensional region a plane through the center which contains two D's *separated*

[5] R. W. Stokes, *A geometric theory of solution of linear inequalities*, Transactions of the American Mathematical Society, vol. 33(1931), p. 804.

[6] "Every point of the convex extension of a closed bounded set lies in the interior or on the boundary of a simplex, the vertices of which belong to the set." (Cf. Bonnesen-Fenchel, p. 9.)

[7] In general one may expect $s' + t' = 7$, since for $s' + t' \leqq 6$ the representative points lie in a five-flat which would have to contain the origin; this requirement would be equivalent to a quadratic or linear condition imposed on y^s and z^t, which is probably not satisfied in general. We shall prove incidentally that $D/\Delta \leqq \sqrt{2}$ for $s' + t' \leqq 5$. An example of a minimal region having actually only 2 greatest diameters was constructed by the author.

by at least two Δ's, in the sense of F. Behrend. It is clear that such a region satisfies the condition of non-separation and is, therefore, a minimal region, if the sufficiency of this condition, which is proved later, is admitted. For this case, F. Behrend has shown that $D/\Delta \leqq \sqrt{2}$.

If the region R has three D's which are coplanar, we conclude in the same manner as in the preceding paragraph that $D/\Delta \leqq \sqrt{2}$ for a minimal region.

It remains to consider only the case where there are three D's not in the same plane and four or fewer Δ's, satisfying the non-separation condition. It must be true of every conic passing through the traces of the D's that it cannot leave all the Δ's on one side. Let any conic through the D's be

$$a_1xy + a_2xz + a_3yz = 0.$$

The condition of non-separation is, then, that the system of inequalities

$$a_1x_iy_i + a_2x_iz_i + a_3y_iz_i > 0 \qquad (i = 1, \cdots, t')$$

must not hold for the coördinates (x_i, y_i, z_i) of the traces of the t' Δ's, and for any a_1, a_2, a_3. That is, for fixed (x_i, y_i, z_i), there must exist non-negative numbers, λ_i, at least one of the λ's being positive, such that

$$\sum \lambda_i(a_1x_iy_i + a_2x_iz_i + a_3y_iz_i) \equiv 0$$

for all a_1, a_2, a_3.[8] The coefficients of the a's must, therefore, vanish:

$$\sum \lambda_ix_iy_i = 0, \quad \sum \lambda_ix_iz_i = 0, \quad \sum \lambda_iy_iz_i = 0. \tag{4}$$

5. **Considerations in three dimensions.** Without restriction of generality for our considerations we may suppose three non-coplanar D's to be two units in length. The convex extension in three-dimensional space of these diameters is an octahedron. The end points of the Δ's must all extend to or through the faces of this octahedron. It is to be proved that the sphere of radius $\frac{1}{3}\sqrt{3}$ does not contain all the end points of the Δ's as interior points.

We consider oblique space coördinates with origin at the center of the convex region and the three axes coinciding with the three D's. The general form for the equation of a sphere with center at the origin is

$$A(x^2 + y^2 + z^2) + Dxy + Exz + Fyz = 1,$$

and the sphere having as its intercepts on the three axes $(\frac{1}{3}\sqrt{3}, 0, 0)$, $(0, \frac{1}{3}\sqrt{3}, 0)$, $(0, 0, \frac{1}{3}\sqrt{3})$ has the equation

$$3(x^2 + y^2 + z^2) + Dxy + Exz + Fyz = 1.$$

[8] W. B. Carver, *Systems of linear inequalities*, Annals of Mathematics, vol. 23(1921–22), p. 217. This follows easily from the fact that the origin lies in the interior of the convex extension of the representative points, as the origin can then be represented as the center of mass of non-negative masses in these points.

Again, let the coördinates of the end points of the Δ's be (x_i, y_i, z_i). Assuming these points are all interior to the sphere gives a proof by contradiction. It is assumed, then, that

$$3(x_i^2 + y_i^2 + z_i^2) + Dx_iy_i + Ex_iz_i + Fy_iz_i < 1.$$

The equations of the eight planes through the end points of the D's are

$$\pm x \pm y \pm z = 1.$$

The Δ's must extend to or through these planes, so that for every (x_i, y_i, z_i) there are some of the above planes such that

$$\pm x_i \pm y_i \pm z_i \geqq 1,$$

or $(\pm x_i \pm y_i \pm z_i)^2 \geqq 1$, for all (x_i, y_i, z_i). If now these equations are subtracted from those relating to the sphere, we get

$$3(x_i^2 + y_i^2 + z_i^2) - (\pm x_i \pm y_i \pm z_i)^2 + Dx_iy_i + Ex_iz_i + Fy_iz_i < 0,$$

where $i = 1, \cdots, t'$ ranges over all the Δ's. Multiply these equations successively by λ_i and add. By the set of equations (4), there results

$$\sum \lambda_i [3(x_i^2 + y_i^2 + z_i^2) - (\pm x_i \pm y_i \pm z_i)^2] < 0,$$

at least one of the λ's being positive and none negative. Schwarz' inequality gives the relation

$$3(x^2 + y^2 + z^2) - (\pm x \pm y \pm z)^2 \geqq 0.$$

From the contradiction, we conclude that the sphere of radius $\frac{1}{3}\sqrt{3}$ does not contain all the end points of the Δ's as interior points. Their end points must either be on this sphere or extend beyond it, and, consequently, $D/\Delta \leqq \sqrt{3}$.

We conclude that in a minimal region with center containing three greatest diameters and any number of Δ's equal to or less than four, the value of $D/\Delta \leqq \sqrt{3}$. In particular, we conclude that in a minimal region with center containing just three D's and four Δ's, the value of $D/\Delta \leqq \sqrt{3}$. The process of polar reciprocation with respect to the unit sphere enables us to make the same statement for a minimal region containing just four D's and three Δ's. If we cause the center of the convex region to coincide with the center of the sphere, it is clear that the relative direction of the D's and Δ's in R', the polar reciprocal figure, will be preserved, and consequently, the condition of non-separation will still hold. Moreover, with every greatest (smallest) diameter of R there will be associated a smallest (greatest) diameter of R'. Also, the ratio of D/Δ will be preserved.

We conclude finally that for convex regions having a center, $\max_{c} \min_{l \subset c} D/\Delta \leqq \sqrt{3}$.

6. **Convex regions without center.** In a convex region without center the greatest and smallest diameters do not necessarily go through one point, and there is no point with respect to which the region is symmetric. The geometric

processes we have used, therefore, become meaningless, for example, the definition of radius no longer holds, there being no central point. By the process of central symmetrization[9] (Zentralsymmetrisierung), however, a region without center may always be converted into a region with center, and, moreover, the magnitude of D and Δ is not changed. The transformed region, R', consists of the points (x', y', z') of the form

$$x' = \tfrac{1}{2}(x_1 - x_2), \qquad y' = \tfrac{1}{2}(y_1 - y_2), \qquad z' = \tfrac{1}{2}(z_1 - z_2),$$

where (x_1, y_1, z_1), (x_2, y_2, z_2) range independently over the points of R. This is clearly an affine-covariant construction, i.e., if an affine transformation carries R into R_1, we have

$$(R_1)' = (R')_1,$$

where the primes denote the process of central symmetrization. Hence, with any class C of convex regions there is associated by the process of central symmetrization a class C', each of whose members has a center and has the same value for D and Δ as class C. It follows that the relation $\max_{c} \min_{l \subset c} D/\Delta \leqq \sqrt{3}$ holds for any convex region.

THEOREM 2. *If l is any closed and bounded convex region in three-dimensional space and C is the class of convex regions equivalent under affine transformations, then* $\max_{c} \min_{l \subset c} D/\Delta \leqq \sqrt{3}$.

We are next led to inquire for what class or classes of convex regions the equality sign in the expression of Theorem 2 holds. From analogy with the case in two dimensions, we should guess that $\max_{c} \min_{l \subset c} D/\Delta = \sqrt{3}$, for the class of parallelepipeds, and the value of D/Δ for the cube is, in fact, $\sqrt{3}$ as one readily verifies. We note that the cube has four D's and three Δ's. This observation leads us to remark that the matter of the number of D's and Δ's throughout our discussion is dual in the sense that where we have counted D's we could just as well have counted Δ's. This fact is implicit in the process of polar reciprocation with respect to the unit sphere applied to any convex region having a center. (See §5.) We can conclude that besides the cube there also exists a minimal region having four Δ's and three D's and with $D/\Delta = \sqrt{3}$. In this case the region is a regular octahedron, this being the polar reciprocal figure to the cube. It is not difficult to prove by elementary methods that a minimal region with just three D's having $D/\Delta = \sqrt{3}$ is necessarily the regular octahedron.

THEOREM 3. *For regions having a center the equality sign of Theorem 2 holds only for the two classes, the parallelepipeds and the octahedrons.*

7. Sufficiency of the condition of non-separation. Let E_1 be an ellipsoid which is circumscribed about the region l, and let E_2 be an ellipsoid which is inscribed in l, where E_1 and E_2 are homothetic with common center at the

[9] Bonnesen-Fenchel, loc. cit., p. 73.

center of l. We denote the affine invariant volume ratio of this *ellipsoid-pair* by $V(E_1, E_2)$. Following F. Behrend, we call an ellipsoid-pair whose ratio is less than or equal to the ratio of all other ellipsoid-pairs of the class C a *minimal ellipsoid-pair.*

Associated with any region l having a center is a minimum circumscribed sphere, $S_D(l)$, of diameter D, and a maximum inscribed sphere, $S_\Delta(l)$, of diameter Δ, the center of both spheres being at the center of l. It is evident that the ratio of this *sphere-pair* for any region l of the class C can not be less than the ratio for l_1, the minimal region of the class, i.e., $V(S_D(l), S_\Delta(l)) \geqq V(S_D(l_1), S_\Delta(l_1))$. We now prove

THEOREM 4. *A region l with center is a minimal region with respect to D/Δ if and only if the sphere-pair associated with it is a minimal ellipsoid-pair.*

Proof. Let us apply to l an affine transformation which carries l into l' and carries any ellipsoid-pair of l into a pair of spheres of l' (not necessarily the sphere-pair). If $V(E_1')$, $V(E_2')$ are, respectively, the volumes of these spheres, we know, of course, that $V(E_1') \geqq V(S_D(l'))$ and $V(E_2') \leqq V(S_\Delta(l'))$, where $V(S_D(l'))$, $V(S_\Delta(l'))$ are the volumes of the sphere-pair of l'. Hence

$$V(E_1, E_2) = V(E_1', E_2') \geqq V(S_D(l'), S_\Delta(l')) \geqq V(S_D(l_1), S_\Delta(l_1)),$$

where l' may or may not be a minimal region l_1. It follows that the sphere-pair of a minimal region is a minimal ellipsoid-pair. The converse is also true. If we have a region l' for which the sphere-pair is a minimum, then

$$\frac{D^3(l')}{\Delta^3(l')} = V(S_D(l'), S_\Delta(l')) \leqq V(S_D(l), S_\Delta(l)) = \frac{D^3(l)}{\Delta^3(l)}$$

for all l of the class C; i.e., l' is a minimal region.

THEOREM 5. *If a region l satisfies the condition of non-separation, the corresponding sphere-pair is a minimal ellipsoid-pair.*

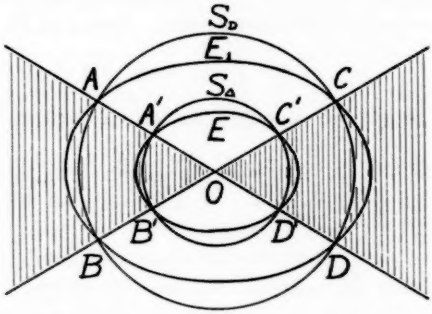
If we prove this, then, by Theorem 4, the converse of Theorem 1 is true; that is, every region satisfying the non-separation condition is a minimal region.

Proof. We assume that the region l satisfies the condition of non-separation. Let E_1 and E_2 be any ellipsoid-pair not both coincident with S_D, S_Δ, respectively, the sphere-pair of l. Theorem 5 is represented by the inequality

$$V(S_D, S_\Delta) \leqq V(E_1, E_2). \tag{5}$$

If E_1 and E_2 are a pair of spheres, this is certainly true. Suppose that E_1 and E_2 are not spheres. We can assume that E_1 intersects the sphere S_D in two conic sections viewed projectively from O, the common center of the ellipsoid-pair and of l. (See the figure.) Both of these sections are symmetric with respect to O; i.e., the point O is the vertex of a cone determined by either of these sections. Let $\overline{AB}$ and $\overline{CD}$ represent these conic sections. The figure represents a cross-section view. Let the corresponding intersections of the

cone on the sphere S_Δ be $\overline{A'B'}$ and $\overline{C'D'}$. Let E be an ellipsoid similar to E_1 and passing through the latter sections.



We now show that E_2 must coincide with or be interior to E. For suppose this were not the case and that E_2 is exterior to E. Any diameter of the convex region, which passes through O and is in the shaded region or on the surface of the cone, must extend beyond E and is greater than Δ. Therefore, the Δ's are restricted to the exterior of the region defined by the cone. Also, this exterior (unshaded) region can contain no D's, since any D cannot extend beyond E_1 and in this region E_1 lies closer to O than S_D. Therefore, the D's are restricted to the interior of the cone or its surface. The condition of non-separation fails to be satisfied, and E_2 is not exterior to E. Since $V(E_1, E) = V(S_D, S_\Delta)$, the relation of equation (5) follows, and the proof of Theorem 5 is complete.

The author desires to acknowledge indebtedness to Dr. Fritz John under whose direction this work was done.

UNIVERSITY OF KENTUCKY.

THE ULTRAHYPERBOLIC DIFFERENTIAL EQUATION WITH FOUR INDEPENDENT VARIABLES

By Fritz John

The properties of a linear homogeneous partial differential equation of the second order

$$(1) \qquad \sum_{i,k} a_{ik} \frac{\partial^2 u}{\partial x_i \partial x_k} + \sum_i b_i \frac{\partial u}{\partial x_i} + cu = 0$$

for a function $u(x_1, \cdots, x_n)$ are known to depend largely on the index of the quadratic form $Q(\xi) = \sum_{i,k} a_{ik} \xi_i \xi_k$.[1] If by a suitable real linear transformation Q can be brought into the form $\pm(\xi_1^2 + \cdots + \xi_n^2)$, i.e., if Q is definite, (1) is called an *elliptic* equation. If Q can be transformed into $\pm(\xi_1^2 + \cdots + \xi_{n-1}^2 - \xi_n^2)$, the equation is called *normal hyperbolic*. Elliptic and normal hyperbolic equations constitute the two types which have been studied more extensively, besides the case of a parabolic equation for which det $(a_{ik}) = 0$. Equations which are neither elliptic, nor parabolic, nor normal hyperbolic, i.e., equations for which the corresponding quadratic form Q can be written in the form $\xi_1^2 + \xi_2^2 \pm \cdots \pm \xi_{n-2}^2 - \xi_{n-1}^2 - \xi_n^2$, have scarcely been treated, at least not without restriction to solutions which are analytic in all or some of the variables. For such equations the notation *ultrahyperbolic* has been introduced by R. Courant.

Ultrahyperbolic equations occur only if the number of independent variables is at least 4. The simplest example is the equation

$$(2) \qquad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} - \frac{\partial^2 u}{\partial x_3^2} - \frac{\partial^2 u}{\partial x_4^2} = 0,$$

which forms the subject of the present paper. Obviously every ultrahyperbolic equation $\sum_{i,k} a_{ik} \frac{\partial^2 u}{\partial x_i \partial x_k} = 0$ with constant coefficients can be transformed into (2).

The theory of ultrahyperbolic equations with constant coefficients has been made accessible recently by the discovery by L. Asgeirsson of a functional equation for the solutions of any second order differential equation with constant coefficients of any type whatsoever.[2] For our equation (2) this functional equation takes the form

Received November 13, 1937.

[1] Cf. Hadamard, *Lectures on Cauchy's Problem*, Book I, Chapter II.

[2] Cf. Asgeirsson, *Über eine Mittelwertseigenschaft von Lösungen homogener linearer partieller Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten*, Mathematische Annalen, vol. 113(1936), pp. 321–346. An exposition of Asgeirsson's results and further applications can be found in the second volume of Courant-Hilbert, *Methoden der Mathematischen Physik*. Cf. also H. Poritsky, *Generalizations of the Gauss law of the spherical mean*, Transactions of the American Mathematical Society, vol. 43(1938), p. 215.

$$\int_0^{2\pi} u(a + r\cos\varphi, b + r\sin\varphi, c, d)\,d\varphi = \int_0^{2\pi} u(a, b, c + r\cos\varphi, d + r\sin\varphi)\,d\varphi \tag{3}$$

and is valid for all a, b, c, d and $r > 0$, if $u(x_1, x_2, x_3, x_4)$ is a twice continuously differentiable solution of (2) in the region

$$[(x_1 - a)^2 + (x_2 - b)^2]^{\frac{1}{2}} + [(x_3 - c)^2 + (x_4 - d)^2]^{\frac{1}{2}} \leqq r.$$

A more general identity is obtained by Asgeirsson from (3) by applying an arbitrary linear transformation which leaves the form $\xi_1^2 + \xi_2^2 - \xi_3^2 - \xi_4^2$ and with it the equation (2) invariant.

In the case of an ultrahyperbolic equation with constant coefficients, it is in general impossible to prescribe the values of a solution on an $(n - 1)$-dimensional manifold. It is proved, e.g., in Courant-Hilbert (loc. cit.), that the values of a solution u of (2) on every non-characteristic $(n - 1)$-dimensional plane region are not independent of one another "im Kleinen" and can be continued only in a unique manner, similarly to analytic functions.[3] There is still the possibility of prescribing the values of u on a characteristic manifold; it will be shown here for the special equation (2) that this is indeed possible for certain characteristic manifolds.

It is the purpose of the present paper to determine the general solution of (2). For this purpose we interpret the independent variables x_1, x_2, x_3, x_4 in a suitable manner as coördinates of a straight line in 3-dimensional xyz-space. u becomes a function of straight lines. Asgeirsson's relation (3) then takes the following simple form: *If the line function u is a solution of* (2), *then for every hyperboloid H of revolution and of one sheet the mean values of u for the two families of generating lines of H are equal.* We call a line function u with the latter property *harmonic*. If u is twice continuously differentiable, harmonic character of u is equivalent to equation (2). The notion of harmonic line function thus represents a generalization of the differential equation (2) to continuous functions, which are not necessarily differentiable. It is more convenient to determine the most general harmonic line function, thus avoiding the difficulties inherent in a differentiability proof.

It is easily seen that the line integrals of a (sufficiently regular) point function $f(x, y, z)$ or the line mean values of a continuous plane function form a harmonic line function. We shall prove that, on the other hand, the most general harmonic line function u (satisfying certain conditions at infinity) can be represented and is uniquely determined by an arbitrary plane function in 3-space. If certain further regularity conditions are satisfied, u can also be represented

[3] In spite of this, the solutions are not necessarily analytic functions. Incidentally we meet with the same behavior in the case of solutions of normal hyperbolic equations on a time-like manifold. Cf. the author's discussion of the solutions of Darboux's equation on a time-like manifold, Mathematische Annalen, vol. 111, p. 549 et seq.

as the line integral of a point function $f(x, y, z)$. With the help of the first representation (which involves fewer assumptions for u), we can solve certain characteristic boundary value problems for equation (2) by reducing these problems to the problem solved by J. Radon, namely, the problem of determining a function in the plane from its line integrals. One can prescribe u on certain characteristic 3-dimensional manifolds in $x_1x_2x_3x_4$-space in such a way that a solution taking these prescribed values exists and is uniquely determined.[4]

1. **Interpretation of u as a line function.** A line in 3-dimensional space may be determined either by two of its points $\xi = (\xi_1, \xi_2, \xi_3)$, $\eta = (\eta_1, \eta_2, \eta_3)$ or by its Plücker coördinates, which we write as follows:

$$\tag{4} \begin{cases} p_1 = \xi_2\eta_3 - \xi_3\eta_2, & q_1 = \xi_1 - \eta_1, \\ p_2 = \xi_3\eta_1 - \xi_1\eta_3, & q_2 = \xi_2 - \eta_2, \\ p_3 = \xi_1\eta_2 - \xi_2\eta_1, & q_3 = \xi_3 - \eta_3. \end{cases}$$

Every ratio $p_1:p_2:p_3:q_1:q_2:q_3$ determines uniquely a line, provided the condition

$$\tag{5} p_1q_1 + p_2q_2 + p_3q_3 = 0$$

is satisfied.[5] We shall write $|\xi|$ for $\sum_i \xi_i^2$ and make use of the summation convention. We assume besides in this section that all functions are twice continuously differentiable in the domains concerned.

THEOREM 1.1. *Let $u(x_1, x_2, x_3, x_4)$ be a solution of* (2) *and let the function v be defined by*

$$\tag{6} \begin{aligned} v(\xi, \eta) &= v(\xi_1, \xi_2, \xi_3, \eta_1, \eta_2, \eta_3) \\ &= \left[\sum_i \left(\frac{q_i}{q_3}\right)^2\right]^{\frac{1}{2}} u\left(\frac{p_2 + q_2}{q_3}, \frac{-p_1 - q_1}{q_3}, \frac{p_2 - q_2}{q_3}, \frac{-p_1 + q_1}{q_3}\right), \end{aligned}$$

where the p_i and q_i denote the expressions given by (4). *Then v is a function of the straight line through ξ and η alone and satisfies the equations*

$$\tag{7} \left(\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial^2}{\partial\xi_k\,\partial\eta_i}\right)\frac{v(\xi, \eta)}{|\xi - \eta|} = 0$$

[4] An equation equivalent to (2) has been treated by G. Hamel in his dissertation (Göttingen, 1901) in connection with the problem of finding all geometries in which the straight lines are the shortest ones (Mathematische Annalen, vol. 57, p. 231 et seq.). There the general solution is determined by certain boundary conditions, but only under the assumption that it is analytic in 2 of its arguments and the results are only proved for a sufficiently small neighborhood. Results similar to those of Hamel for more general differential equations have been obtained recently by E. W. Titt, but are still unpublished. Cf. the abstract in the Bull. Am. Math. Soc., vol. 42(1936), p. 32.

[5] For the elementary properties of line coördinates cf. F. Klein, *Höhere Geometrie*, §20 et seq. The deeper connection between the ultrahyperbolic differential equation (2) and the theory of straight lines in space is given by the identity (5).

for all i, k = 1, 2, 3. (There are essentially 3 such equations.) If on the other hand $v(\xi, \eta)$ depends on the straight line through ξ and η alone and satisfies the equations (7), then v can be represented in the form (6) with the help of a solution u of (2).

Proof. If v is defined by (6), it obviously depends only on the line through ξ and η. Equation (7) can be easily verified for $i = 1$, $k = 2$; from that it can be derived for any i, k in the following way. The fact that v is a line function implies that for every ϑ

$$v(\vartheta\xi + (1 - \vartheta)\eta, \eta) = v(\xi, \vartheta\eta + (1 - \vartheta)\xi) = v(\xi, \eta). \tag{8}$$

Thus also

$$\frac{v(\vartheta\xi + (1 - \vartheta)\eta, \eta)}{|\xi + (1 - \vartheta)\eta - \eta|} = \frac{1}{\vartheta}\frac{v(\xi, \eta)}{|\xi - \eta|}.$$

On differentiating with respect to ϑ and putting $\vartheta = 1$, we obtain

$$(\xi_i - \eta_i)\frac{\partial}{\partial\xi_i}\frac{v(\xi, \eta)}{|\xi - \eta|} + \frac{v(\xi, \eta)}{|\xi - \eta|} = 0.$$

If we differentiate with respect to η_k, it follows that

$$\left[(\xi_i - \eta_i)\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial}{\partial\xi_k} + \frac{\partial}{\partial\eta_k}\right]\frac{v(\xi, \eta)}{|\xi - \eta|} = 0.$$

If we add to this equation the one obtained by interchanging ξ and η, we get

$$(\xi_i - \eta_i)\left(\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial^2}{\partial\xi_k\,\partial\eta_i}\right)\frac{v(\xi, \eta)}{|\xi - \eta|} = 0$$

for $k = 1, 2, 3$. Hence all relations (7) follow if (7) is already verified for $i = 1$, $k = 2$. (Here it is essential that i and k are restricted to the values 1, 2, 3.)

If on the other hand v is a line function satisfying (7), v can be represented in the form (6) with the help of some function u, since v depends only on $p_1:p_2:q_1:q_2:q_3$ (the ratio to p_3 is determined by (5)). It follows for $\xi_3 = 0$, $\eta_3 = 1$ that

$$\frac{v(\xi_1, \xi_2, 0, \eta_1, \eta_2, 1)}{[(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + 1]^{\frac{1}{2}}} = u(\xi_1 - \xi_2 + \eta_2, \xi_1 + \xi_2 - \eta_1, \xi_1 + \xi_2 - \eta_2, -\xi_1 + \xi_2 + \eta_1);$$

from this relation (2) can be easily derived by using (7) for $i = 1$, $k = 2$. This completes the proof of the theorem.

On the basis of Theorem 1.1 we can replace the equation (2) for a function $u(x_1, x_2, x_3, x_4)$ by the system of equations (7) for a line function $v(\xi, \eta)$, which differs only by a simple factor from u. The line with the Plücker coördinates $p_1, p_2, \cdots, q_3$ in xyz-space corresponds to the point with coördinates

$$x_1 = \frac{p_2 + q_2}{q_3}, \quad x_2 = \frac{-p_1 - q_1}{q_3}, \quad x_3 = \frac{p_2 - q_2}{q_3}, \quad x_4 = \frac{-p_1 + q_1}{q_3} \tag{9}$$

in $x_1x_2x_3x_4$-space. Conversely, we have, using (5),

$$(10)\quad p_1:p_2:p_3:q_1:q_2:q_3 = \tfrac{1}{2}(-x_2 - x_4):\tfrac{1}{2}(x_1 + x_3):\tfrac{1}{4}(-x_1^2 - x_2^2 + x_3^2 + x_4^2)$$
$$:\tfrac{1}{2}(-x_2 + x_4):\tfrac{1}{2}(x_1 - x_3):1.$$

But the following facts are to be observed. The assumption that v is defined and twice continuously differentiable also for the lines parallel to the xy-plane ($q_3 = 0$), imposes certain conditions on the behavior of u at infinity. On the other hand, if the lines are given by two points ξ and η, v will be undetermined for $\xi = \eta$, but will have a limit, if η approaches ξ from a definite direction; this latter singularity is of course only due to the special representation of the lines. Besides we agree always to choose the positive value for the square root in (6).

THEOREM 1.2. *The equations* (7) *are invariant under simultaneous similarity transformations of* ξ *and* η. *More generally, if the line function* $v(\xi, \eta)$ *satisfies* (7), *then for arbitrary* a_{ik}, b_i *with* $\det(a_{ik}) \neq 0$, *the line function* w *defined by*

$$w(\xi, \eta) = \frac{\sum_i (\xi_i - \eta_i)^2}{\sum_i [a_{ik}(\xi_k - \eta_k)]^2}\, v(a_{rs}\xi_s + b_r, c_{rs}\eta_s + b_r)$$

satisfies (7). (*If the linear transformation defined by the* a *and* b *is orthogonal,* w *reduces to* $v(a_{ik}\xi_k + b_i, a_{ik}\eta_k + b_i)$.)

Proof. That w is again a function of the lines through ξ and η alone is obvious. Moreover, if we write for the moment

$$\sigma(\xi, \eta) = \frac{v(\xi, \eta)}{|\xi - \eta|},$$

it follows from (7) that

$$\left(\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial^2}{\partial\xi_k\,\partial\eta_i}\right)\sigma(\xi, \eta) = 0.$$

Then also

$$\begin{aligned}\left(\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial^2}{\partial\xi_k\,\partial\eta_i}\right)\frac{w(\xi, \eta)}{|\xi - \eta|} &= \left(\frac{\partial^2}{\partial\xi_i\,\partial\eta_k} - \frac{\partial^2}{\partial\xi_k\,\partial\eta_i}\right)\sigma(a_{rs}\xi_s + b_r, a_{rs}\eta_s + b_r)\\ &= a_{ri}a_{sk}\frac{\partial^2\sigma}{\partial\xi_r\,\partial\eta_s} - a_{rk}a_{si}\frac{\partial^2\sigma}{\partial\xi_r\,\partial\eta_s}\\ &= a_{ri}a_{sk}\frac{\partial^2\sigma}{\partial\xi_r\,\partial\eta_s} - c_{sk}a_{ri}\frac{\partial^2\sigma}{\partial\xi_s\,\partial\eta_r}\\ &= a_{sk}a_{ri}\left(\frac{\partial^2}{\partial\xi_r\,\partial\eta_s} - \frac{\partial^2}{\partial\xi_s\,\partial\eta_r}\right)\sigma = 0.\end{aligned}$$

2. **Asgeirsson's theorem in terms of line functions.** We now apply the relation (3) of Asgeirsson to equation (2) in the simplest case $a = b = c = d = 0$, assuming that u is a twice continuously differentiable solution of (2) for

$$(x_1^2 + x_2^2)^{\frac{1}{2}} + (x_3^2 + x_4^2)^{\frac{1}{2}} \leqq r.$$

Translating this formula for u into one for v with the help of (6), we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} v\left(-\tfrac{1}{2}r\cos\varphi,\ -\tfrac{1}{2}r\sin\varphi,\ 0,\ \frac{-r\cos\varphi + r\sin\varphi}{2},\ \frac{-r\cos\varphi - r\sin\varphi}{2},\ 1\right)d\varphi$$

$$= \frac{1}{2\pi}\int_0^{2\pi} v\left(-\tfrac{1}{2}r\cos\varphi,\ -\tfrac{1}{2}r\sin\varphi,\ 0,\ \frac{-r\cos\varphi - r\sin\varphi}{2},\ \frac{r\cos\varphi - r\sin\varphi}{2},\ 1\right)d\varphi.$$

The lines forming the arguments of v in these two integrals for varying φ are respectively the two families of generating lines of the hyperboloid H_0 of revolution of one sheet

$$x^2 + y^2 = \tfrac{1}{4}r^2(1 + z^2), \tag{11}$$

and $d\varphi$ can be interpreted as the differential of the angle POQ, where O is the center of H_0, P the point of intersection of a generating line with the equatorial plane (i.e., the plane perpendicular to the axis of revolution through O), and Q some fixed point in that plane. We shall prove below that the region

$$(x_1^2 + x_2^2)^{\frac{1}{2}} + (x_3^2 + x_4^2)^{\frac{1}{2}} \leqq r$$

corresponds to the lines in the interior of H_0. Thus, if v is a twice continuously differentiable line function satisfying (7) for all lines in the interior and on the boundary of the hyperboloid H_0, then the mean values of v over the two families of generating lines of H_0 are equal.

As every hyperboloid of revolution of one sheet is similar to some hyperboloid of the form (11), we have according to Theorem 1.2

THEOREM 2.1. *If the line function v is a twice continuously differentiable solution of* (7) *for all lines in the interior and on the boundary of a hyperboloid of revolution of one sheet H, then the mean values of v over the two families of generating lines of H are equal* (*the variable of integration φ is chosen as above*).

If we apply to H_0 an arbitrary affine transformation we obtain a mean value theorem involving the generating lines of an arbitrary hyperboloid of one sheet (not necessarily of revolution).

THEOREM 2.2. *If the line function v is a twice continuously differentiable solution of* (7) *for all lines in the interior and on the boundary of a hyperboloid of one sheet H, then the mean values of v* cos γ *are the same for both families of generating lines of H. Here γ denotes the angle the argument line of v makes with the principal non-intersecting axis of H; the variable of integration φ is the same as before.* (*For a hyperboloid of revolution γ = const.*)

We pass over the elementary calculations leading to a proof of Theorem 2.2 along the lines indicated above, as no use will be made of this theorem in the following paragraphs.

We still have to prove that the lines in the interior of the hyperboloid (11) in xyz-space correspond to the points with $(x_1^2 + x_2^2)^{\frac{1}{2}} + (x_3^2 + x_4^2)^{\frac{1}{2}} < r$. For this purpose we notice that the condition that two lines with line coördinates $(p_1, p_2, \cdots, q_3)$ and $(p_1', p_2', \cdots, q_3')$ intersect is

$$p_1q_1' + p_2q_2' + p_3q_3' + q_1p_1' + q_2p_2' + q_3p_3' = 0$$

(cf. Klein, loc. cit., p. 84). Then, by using (9), the condition that two points (x_1, x_2, x_3, x_4) and (x_1', x_2', x_3', x_4') in $x_1x_2x_3x_4$-space correspond to two intersecting lines in xyz-space may be written:

$$Q(x - x') = (x_1 - x_1')^2 + (x_2 - x_2')^2 - (x_3 - x_3')^2 - (x_4 - x_4')^2 = 0.$$

We call two points in $x_1x_2x_3x_4$-space for which this relation holds "incident". The line joining two incident points is a characteristic of the differential equation (2).

The lines not intersecting the hyperboloid (11) are those not intersecting any of the generating lines of one of the families of H_0. One of these families is represented by the circle $x_1^2 + x_2^2 = r^2$, $x_3 = x_4 = 0$ or

$$(12) \qquad x_1 = r \cos \varphi, \qquad x_2 = r \sin \varphi, \qquad x_3 = x_4 = 0.$$

The lines not intersecting H_0 correspond to points which are incident with no point of the form (12), i.e., points for which the equation

$$(x_1 - r \cos \varphi)^2 + (x_2 - r \sin \varphi)^2 - x_3^2 - x_4^2 = 0$$

has no real solutions φ. This is the case if either

$$|(x_3^2 + x_4^2)^{\frac{1}{2}} - (x_1^2 + x_2^2)^{\frac{1}{2}}| > r \qquad \text{or} \qquad (x_1^2 + x_2^2)^{\frac{1}{2}} + (x_3^2 + x_4^2)^{\frac{1}{2}} < r.$$

The point sets corresponding to these two cases have no common points. Hence it is the second case that corresponds to the non-intersecting lines in the *interior* of H_0, e.g., the z-axis, which is an interior line, corresponds to $x_1 = x_2 = x_3 = x_4 = 0$ and satisfies the second inequality. The points satisfying the first inequality must thus correspond to the lines exterior to H_0.

It can be shown that Theorem 2.2 contains Asgeirsson's theorem for equation (2) in its most general form, obtained by applying to (3) an arbitrary linear transformation which leaves $x_1^2 + x_2^2 - x_3^2 - x_4^2$ invariant. For this purpose one has to find the configuration of the points in $x_1x_2x_3x_4$-space corresponding to the generating lines of an arbitrary hyperboloid H of one sheet. If we start with the definition of a hyperboloid as the locus of the lines intersecting 3 given lines, it is easy to prove that the two families of generating lines of H correspond to a pair of "conjugate conics" in $x_1x_2x_3x_4$-space, i.e., a pair of conics which can be generated as follows. Take any point x^0 and consider the characteristic cone $Q(x - x^0) = 0$ with vertex x^0. Take a pair of 2-dimensional planes π_1 and π_2 through x^0 which are conjugate to this cone. A pair of conjugate conics is obtained by cutting π_1 and π_2 with the two surfaces $Q(x - x^0) = c$ and $Q(x - x^0) = -c$, respectively, c being an arbitrary constant. If H contains lines parallel to the xy-plane, then the corresponding conjugate conics are hyperbolas, otherwise ellipses.

Our Theorem 2.2 states that the integrals of a solution of (2) over a pair of conjugate conics are equal if a certain variable of integration is used. This statement is equivalent to the theorem of Asgeirsson in the most general form if the conics are ellipses. The case of conjugate hyperbolas, on the other hand,

is new; it depends of course on the special conditions for u at infinity, which correspond to the assumption that v is still regular for lines parallel to the xy-plane. This integral formula for conjugate hyperbolas could not be derived like the case of conjugate ellipses by linear transformations in $x_1x_2x_3x_4$-space from the theorem in its special form (3) involving conjugate circles, as, of course, circles could not be transformed into hyperbolas by an affine transformation. Indeed the affine transformations of xyz-space do not induce linear transformations of $x_1x_2x_3x_4$-space, but linear transformations on the expressions p_1, p_2, p_3, q_1, q_2, q_3 given by (10), which leave the relation (5) invariant.

A 2-dimensional plane π is called *characteristic*, if it is conjugate to itself. In this case for every x and x^0 on π

$$Q(x - x^0) = 0;$$

thus π is contained in some characteristic cone. A family of parallel lines in xyz-space, i.e., those lines for which the ratio $q_1:q_2:q_3$ is constant, corresponds to a characteristic two-dimensional plane of a special kind; in this case

$$x_4 - x_2 = \text{const.} = c, \qquad x_1 - x_3 = \text{const.} = d.$$

Thus, the lines parallel to a given line correspond to a characteristic 2-flat which is parallel to the 2-flat $x_4 - x_2 = 0$, $x_1 - x_3 = 0$.

3. Harmonic line functions.

DEFINITION. A function $v = v(l)$ of the lines l of xyz-space is called *harmonic* if v is continuous and if for every hyperboloid of revolution H of one sheet, which is such that v is defined for all lines in the interior and on the boundary of H, the mean values of v over the two families of generating lines of H are equal. (As variable of integration φ naturally the polar angle in the equatorial plane is to be taken.)

Theorem 2.1 may now be stated thus:

THEOREM 3.1. *A line function v which is twice continuously differentiable and satisfies equations* (7) *is harmonic.*

THEOREM 3.2. *Every harmonic line function which is twice continuously differentiable satisfies the equations* (7).

Proof. Consider v for the lines neighboring any given line l. Without restriction of generality we may assume that l is the z-axis, as our properties are invariant under orthogonal transformations. Then v satisfies the mean value theorem for all hyperboloids of revolution about the z-axis of sufficiently small radius. Hence the corresponding function u satisfies the mean value theorem (3) for $a = b = c = d = 0$ and sufficiently small r. From this it follows, according to Asgeirsson, that u satisfies (2) at the origin and consequently v satisfies (7) at z.[6]

Theorems 3.1 and 3.2 permit us to consider harmonic character of a line

[6] Cf. Asgeirsson, loc. cit., p. 336.

function v as a generalization of the differential equations (7) to line functions, which are only continuous. (As was pointed out by Asgeirsson, loc. cit., the mean value property (3) for u or the equivalent harmonic character of v alone would not be sufficient to assure that the functions are twice continuously differentiable. This is in contrast to the analogous case of potential functions, where the mean value theorem alone already implies analyticity of the function.) We shall determine the general harmonic line function rather than the general solution of (7), thus avoiding the difficulties inherent in a differentiability proof. Two classes of harmonic line functions will be constructed in the next two theorems, and we shall prove later that they furnish us the most general harmonic line functions, subject to certain regularity conditions.

THEOREM 3.3. *Let $f(x, y, z)$ be a continuous point function for which there exists a monotonic function $h(r)$ such that*

$$|f(x, y, z)| \leqq h((x^2 + y^2 + z^2)^{\frac{1}{2}})$$

and for which $\int_0^\infty h(r)dr$ converges. Then the integrals of f over the straight lines of xyz-space form a harmonic line function.

Proof. Without restriction of generality we consider the hyperboloid H_0: $x^2 + y^2 = r^2(1 + z^2)$, which has the generating lines

$$x = r(\cos\varphi \pm t \sin\varphi), \qquad y = r(\sin\varphi \mp t\cos\varphi), \qquad z = t.$$

Let $v = v(\varphi)$ denote the line integral of f over the line with parameter φ. Then' if $ds = (r^2 + 1)^{\frac{1}{2}} dt$ is the element of length of arc on the line,

$$\begin{aligned}\int_0^{2\pi} v(\varphi)\, d\varphi &= \int_0^{2\pi} d\varphi \int_{-\infty}^{+\infty} f(x, y, z)\, ds \\ &= (r^2 + 1)^{\frac{1}{2}} \int_0^{2\pi}\int_{-\infty}^{+\infty} f(x, y, z)\, dt\, d\varphi \\ &= (1 + r^{-2})^{\frac{1}{2}} \int\int_{x^2+y^2>r^2} \frac{f(x, y, r^{-1}(x^2 + y^2 - r^2)^{\frac{1}{2}})}{(x^2 + y^2 - r^2)^{\frac{1}{2}}}\, dx\, dy\end{aligned}$$

irrespective of which of the two families of generating lines of H_0 is used. The improper integrals are uniformly convergent. This proves the theorem.

Remark. A more general theorem similar to Theorem 2.2 involving the generating lines of any hyperboloid of one sheet could be derived as well.

THEOREM 3.4. *Let p be a fixed positive number and F be an arbitrary continuous function of the planes in xyz-space. If $v(l)$ denotes the mean value of F over the planes tangent to the cylinder of radius p with axis l, then $v(l)$ is a harmonic line function.*

Proof. It is again sufficient to prove that the mean values of v over the two families of generating lines of the hyperboloid H_0

$$x^2 + y^2 = r^2(1 + z^2)$$

are equal. We use (ξ, η, ζ) as coördinates of the plane $\xi x + \eta y + \zeta z = 1$. The equation of H_0 in plane coördinates is then

$$\xi^2 + \eta^2 = r^{-2}(1 + \zeta^2).$$

The generating lines are given in plane coördinates by

$$\xi = r^{-1}(\cos \varphi \pm t \sin \varphi), \qquad \eta = r^{-1}(\sin \varphi \mp t \cos \varphi), \qquad \zeta = t$$

(upper sign for one family, lower for the other family). For varying t and fixed φ, we obtain all planes through the generating line l_φ with parameter φ. A plane parallel to the plane with parameters φ, t is given by

$$\xi = \sigma^{-1}r^{-1}(\cos \varphi \pm t \sin \varphi), \qquad \eta = \sigma^{-1}r^{-1}(\sin \varphi \mp t \cos \varphi), \qquad \zeta = \sigma^{-1}t;$$

the distance of this plane from the parallel plane through l_φ is $(\sigma - 1)r(1 + t^2 + r^2t^2)^{-\frac{1}{2}}$. If this distance is to be equal to p, we have to choose $\sigma = 1 + pr^{-1}(1 + t^2 + r^2t^2)^{\frac{1}{2}}$. For the angle ϑ between the plane (φ, t) and the plane $(\varphi, 0)$ $\cos \vartheta = (1 + t^2 + r^2t^2)^{-\frac{1}{2}}$. Thus

$$t = (1 + r^2)^{-\frac{1}{2}} \tan \vartheta, \qquad \sigma = 1 + \frac{p}{r \cos \vartheta}.$$

Hence the mean value of the plane function $F(\xi, \eta, \zeta)$ over the cylinder of radius p about the generating line l_φ is given by

$$\begin{aligned} v(l_\varphi) &= \frac{1}{2\pi} \int_0^{2\pi} F(\xi, \eta, \zeta)\, d\vartheta = \frac{1}{2\pi} \int_0^{2\pi} F\left(\frac{\cos \varphi \pm t \sin \varphi}{\sigma r}, \frac{\sin \varphi \mp t \cos \varphi}{\sigma r}, \frac{t}{\sigma}\right) d\vartheta \\ &= \frac{1}{2\pi} \int_0^{2\pi} F\left(\frac{\cos \varphi \cos \vartheta \pm (1 + r^2)^{-\frac{1}{2}} \sin \varphi \sin \vartheta}{\cos \vartheta + pr^{-1}},\right. \\ &\qquad \left.\frac{\sin \varphi \cos \vartheta \mp (1 + r^2)^{-\frac{1}{2}} \cos \varphi \sin \vartheta}{\cos \vartheta + pr^{-1}}, \frac{(1 + r^2)^{-\frac{1}{2}} \sin \vartheta}{\cos \vartheta + pr^{-1}}\right) d\vartheta. \end{aligned}$$

Let $\psi = \psi(\vartheta)$ be defined by

$$\cos \psi = (1 + r^2 \cos^2 \vartheta)^{-\frac{1}{2}}(1 + r^2)^{\frac{1}{2}} \cos \vartheta, \qquad \sin \psi = \pm (1 + r^2 \cos^2 \vartheta)^{-\frac{1}{2}} \sin \vartheta.$$

(The definition of ψ depends on the choice of the family of generating lines.) Then

$$\begin{aligned} v(l_\varphi) = \frac{1}{2\pi} \int_0^{2\pi} F\Big(&(1 + r^2 \cos^2 \vartheta)^{\frac{1}{2}}(1 + r^2)^{-\frac{1}{2}} \cos (\varphi - \psi), \\ &(1 + r^2 \cos^2 \vartheta)^{\frac{1}{2}}(1 + r^2)^{-\frac{1}{2}} \sin (\varphi - \psi), \frac{(1 + r^2)^{-\frac{1}{2}} \sin \vartheta}{\cos \vartheta + pr^{-1}}\Big) d\vartheta. \end{aligned}$$

Consequently the mean value of $v(l)$ over a family of generating lines of H_0 is in easily understood notation

$$\frac{1}{2\pi}\int_0^{2\pi} v(l_\varphi)\,d\varphi = \frac{1}{(2\pi)^2}\int_0^{2\pi} d\vartheta \int_0^{2\pi} F(\varphi, \vartheta)\,d\varphi = \frac{1}{(2\pi)^2}\int_0^{2\pi} d\vartheta \int_0^{2\pi} F(\varphi + \psi(\vartheta), \vartheta)\,d\varphi$$

$$= \frac{1}{(2\pi)^2}\int_0^{2\pi} d\vartheta \int_0^{2\pi} F\Big((1 + r^2\cos^2\vartheta)^{\frac{1}{2}}(1 + r^2)^{-\frac{1}{2}}\cos\varphi,$$

$$(1 + r^2\cos^2\vartheta)^{\frac{1}{2}}(1 + r^2)^{-\frac{1}{2}}\sin\varphi,\ \frac{(1 + r^2)^{-\frac{1}{2}}\sin\vartheta}{\cos\vartheta + pr^{-1}}\Big)\,d\varphi,$$

and this expression now is the same, irrespective for which family of generating lines of H_0 it is formed.

4. The plane integrals of a harmonic line function. We now turn to the consideration of the integral of a harmonic line function v over a family of parallel lines in a plane. We shall prove that, under suitable conditions on the behavior of v for the lines far away from the origin, this integral is the same for every family of parallel lines in the same plane. This will be proved as a limiting case of Theorem 2.1, which is obtained by letting a hyperboloid degenerate into a plane, its generating lines going over into 2 families of parallels in that plane.

DEFINITION. A line function $v(l)$ may be called *regular at infinity* if there exists a monotonic function $h(\Delta)$ for which $\int_0^\infty h(\Delta)d\Delta$ converges and which is such that for every line l for which v is defined $|v(l)| \leqq h(\Delta)$, if Δ is the distance of l from the origin.

THEOREM 4.1. *Let $v(l)$ be a harmonic line function which is regular at infinity and defined and continuous for all lines in 3-space. Then the integrals of v over two families of parallels in the same plane are equal (if the distance of a line from a fixed point in that plane is taken as variable of integration).*

Proof. We apply Theorem 2.1 to the generating lines of the hyperboloid of revolution $H_{\lambda,c}$,

$$(x - \lambda)^2 + y^2 = \left(\frac{z}{c}\right)^2 + \left(\lambda - \frac{z}{c}\right)^2,$$

which are given in parametric representation by

$$x = \lambda(1 - \cos\varphi) + \lambda t(\pm\sin\varphi + \cos\varphi),$$

$$y = -\lambda\sin\varphi + \lambda t(\sin\varphi \mp \cos\varphi),$$

$$z = \lambda ct$$

(the upper sign for one family of generating lines of H, the lower for the other family). Thus for a harmonic function v the two integrals

$$I_\pm = \int_{-\pi}^{+\pi} v(\lambda(1 - \cos\varphi), -\lambda\sin\varphi, 0, \lambda(1 \pm \sin\varphi), \mp\lambda\cos\varphi, \lambda c)\, d\varphi$$

have the same values, whether the upper or lower sign is chosen.

Using the fact that $v(\xi, \eta)$ is dependent only on the line through the points ξ and η, we have (cf. formula (8))

$$v(\xi, \lambda\eta) = v(\xi, \lambda\eta + (1 - \lambda^{-1})(\xi - \lambda\eta)) = v(\xi, \eta + \lambda^{-1}(\lambda - 1)\xi).$$

Accordingly we may write $I_\pm$

$$\begin{aligned}\int_{-\pi}^{+\pi} v(\lambda(1 - \cos\varphi), -\lambda\sin\varphi, 0, 1 \pm \sin\varphi + (\lambda - 1)(1 - \cos\varphi),\\ \mp\cos\varphi - (\lambda - 1)\sin\varphi, c)\, d\varphi\\ = \int_{-\pi}^{+\pi} v(0, -\lambda\varphi, 0, 1, \mp 1 - \lambda\varphi, c)\, d\varphi + \int_{-\pi}^{+\pi} (v(l_1) - v(l_2))\, d\varphi,\end{aligned}$$

where l_1 is the line through the points

$$(\lambda(1 - \cos\varphi), -\lambda\sin\varphi, 0),$$
$$(1 \pm \sin\varphi + (\lambda - 1)(1 - \cos\varphi), \mp\cos\varphi - (\lambda - 1)\sin\varphi, c)$$

and l_2 is the line through

$$(0, -\lambda\varphi, 0), \qquad (1, \mp 1 - \lambda\varphi, c).$$

The distance d between l_1 and l_2 is less than or equal to the distance between any point of l_1 and any point of l_2. Thus

$$d \leqq [\lambda^2(1 - \cos\varphi)^2 + \lambda^2(\varphi - \sin\varphi)^2]^{\frac{1}{2}} \leqq A\lambda\varphi^2,$$

where A is independent of λ and φ. For the angle ψ between l_1 and l_2 we have

$$\sin\psi = \frac{2}{2 + c^2}[\sin^2\varphi + c^2(1 - \cos\varphi)]^{\frac{1}{2}} \leqq A|\varphi|.$$

Finally the distance of l_1 from the origin is

$$\begin{aligned}\Delta_1 &= \lambda\left[\frac{(1 \pm \sin\varphi - \cos\varphi)^2 + 2c^2(1 - \cos\varphi)}{2 + c^2}\right]^{\frac{1}{2}}\\ &\geqq \lambda\left[\frac{2c^2(1 - \cos\varphi)}{2 + c^2}\right]^{\frac{1}{2}} \geqq B\lambda|\varphi|\end{aligned}$$

and that of l_2

$$\Delta_2 = \lambda\varphi\left[\frac{1 + c^2}{2 + c^2}\right]^{\frac{1}{2}} \geqq B\lambda|\varphi|,$$

where B is independent of λ and φ.

That v is uniformly continuous for all lines follows from the continuity of v and the fact that $\lim_{\Delta\to\infty} v = 0$; thus for every ϵ there is a $\delta = \delta(\epsilon)$ such that

$$| v(l_1) - v(l_2) | \leqq \epsilon$$

for $d < \delta$ and $| \sin \psi | < \delta$, and hence for $d + | \sin \psi | < \delta$.

Since $\int_0^\infty h(\Delta)d\Delta$ converges, there is an M such that

$$\int_M^\infty h(B\Delta)\, d\Delta < \epsilon.$$

Let λ be such that

$$\lambda > \frac{AM + AM^2}{\delta(\epsilon M^{-1})} \quad \text{and} \quad \lambda > \frac{M}{\pi}.$$

Then $| v(l_1) - v(l_2) | \leqq \epsilon M^{-1}$ for $d + | \sin \psi | < \lambda^{-1}(AM + AM^2)$, and hence for $| \varphi | < M\lambda^{-1}$. Therefore

$$\begin{aligned}
\left| \lambda \int_{-\pi}^{+\pi} [v(l_1) - v(l_2)]\, d\varphi \right| &\leqq \left| \lambda \int_{-M/\lambda}^{+M/\lambda} [v(l_1) - v(l_2)]\, d\varphi \right| + \left| \lambda \int_{M/\lambda}^{\pi} [v(l_1) - v(l_2)]\, d\varphi \right| \\
&\quad + \left| \lambda \int_{-\pi}^{-M/\lambda} [v(l_1) - v(l_2)]\, d\varphi \right| \\
&\leqq \lambda \cdot 2M\lambda^{-1} \cdot \epsilon M^{-1} + \left| \lambda \int_{-\pi}^{-M/\lambda} 2h(B\lambda | \varphi |)\, d\varphi \right| + \left| \lambda \int_{M/\lambda}^{\pi} 2h(B\lambda | \varphi |)\, d\varphi \right| \\
&\leqq 2\epsilon + 4 \int_M^{\pi\lambda} h(B\varphi)\, d\varphi \leqq 2\epsilon + 4 \int_M^\infty h(B\varphi)\, d\varphi \leqq 6\epsilon.
\end{aligned}$$

Thus

$$\lim_{\lambda\to\infty} \lambda \int_{\pi}^{+\pi} (v(l_1) - v(l_2))\, d\varphi = 0.$$

Consequently

$$\begin{aligned}
\lim_{\lambda\to\infty} \lambda I_\pm &= \lim_{\lambda\to\infty} \lambda \int_{-\pi}^{+\pi} v(0, -\lambda\varphi, 0, 1, \mp 1 - \lambda\varphi, c)\, d\varphi \\
&= \lim_{\lambda\to\infty} \int_{-\pi}^{+\pi} v(0, -\varphi, 0, 1, \mp 1 - \varphi, c)\, d\varphi \\
&= \int_{-\infty}^{+\infty} v(0, -\varphi, 0, 1, \mp 1 - \varphi, c)\, d\varphi.
\end{aligned}$$

The improper integrals exist, as the argument line has the distance $\Delta = \varphi(1 + c^2)^{\frac{1}{2}} (2 + c^2)^{-\frac{1}{2}} > \frac{1}{2}\varphi$ from the origin, and hence $| v | < h(\frac{1}{2}\varphi)$. If we put $x = -\varphi$, we finally obtain the identity

$$\int_{-\infty}^{+\infty} v(0, x, 0, 1, -1 + x, c)\, dx = \int_{-\infty}^{+\infty} v(0, x, 0, 1, +1 + x, c)\, dx$$

for every $c > 0$. This formula may be interpreted as stating that the integrals of v over a pair of families of parallel lines, which include an angle $\omega = \text{arc cos } c^2(2 + c^2)^{-1}$ and lie in the same plane $x = zc^{-1}$, are equal if the distance of a line from a point of the plane is taken as variable of integration. As our assumptions for v are invariant under orthogonal transformations and c is arbitrary, this holds for any two families of parallel lines in one and the same arbitrary plane.

If v satisfies the assumptions of Theorem 4.1, then, for every plane π of xyz-space, a value $F(\pi)$ is uniquely determined, namely, the value of the integral of v over any family of parallels in π. We call $F(\pi)$ the *plane integral* of v. $F(\pi)$ is already uniquely determined for a plane π if v is given for some family of parallels in π.

THEOREM 4.2. *$F(\pi)$ is continuous and $\lim_{p\to\infty} F(\pi) = 0$, if p is the distance of π from the origin.*

Proof. Let π_1 and π_2 be two planes; $F(\pi_1)$ and $F(\pi_2)$ can be expressed as integrals of v over the families of lines parallel to the intersection l_0 of π_1 and π_2. Let M be such that $\int_M^\infty h(\Delta)d\Delta < \epsilon$; let δ be such that $| v(l_1) - v(l_2) | < \epsilon M^{-1}$, if l_1 and l_2 are two parallels of distance $< \delta$. If then π_2 is so near to π_1 that parallels in π_1 and π_2, having the same distance from l_0 and a distance $\Delta \leqq M$ from the origin, have a distance $\leqq \delta$ from one another, then $| F(\pi_1) - F(\pi_2) | \leqq 4\epsilon$, and this proves the first part of our statement. Moreover, obviously, if p is the distance of π from the origin,

$$|F(\pi)| \leqq 2\int_0^\infty h((p^2 + x^2)^{\frac{1}{2}})\,dx = 2\int_0^M h((p^2 + x^2)^{\frac{1}{2}})\,dx + 2\int_M^\infty h((p^2 + x^2)^{\frac{1}{2}})\,dx$$

$$\leqq 2\int_0^M h(p)\,dx + 2\int_M^\infty h(x)\,dx \leqq 4\epsilon,$$

if $M = M(\epsilon)$ is chosen in such a way that $\int_M^\infty h(x)dx < \epsilon$ and p is so large that $h(p) < \epsilon M^{-1}$.

THEOREM 4.3. *If v is a harmonic line function which is regular at infinity, then v is uniquely determined by $F(\pi)$. More exactly, $v(l_0)$ is uniquely determined if $F(\pi)$ is known for all planes π parallel to l_0.*

Proof. Let the lines parallel to l_0 be determined by their points of intersection P with some fixed plane π_0 perpendicular to l_0. We define a function $w(P)$ for all points P in π_0 by $w(P) = v(l)$, where l is the parallel through P to l_0; and a function $\phi(p)$ for all lines p of π_0 by $\phi(p) = F(\pi)$, where π is the plane parallel to l_0 through p. Then $\phi(p)$ is the integral of $w(P)$ over the line p. Besides, there exists a monotonic function $h(r)$ for which $\int_0^\infty h(r)dr$ converges and such that for a point P of π_0 of distance r from the origin $| w(P) | < h(r)$ (the origin shall be the point of intersection of l_0 and π_0). We also know from Theorem 4.2, that $\phi(p)$ is uniformly continuous.

Theorem 4.3 will be proved if we can show that $w(0)$ is uniquely determined by $\phi(p)$, i.e., that the point function w in the plane π_0 is uniquely determined by its line integrals. This was proved by J. Radon (also generalizations to higher dimensions).[7] We reproduce here Radon's simple method in a form adapted to our assumptions.

Let the line p in π_0 be determined by its distance r from O and the polar angle φ of its normal. Then $\phi(p) = \phi(r, \varphi)$. Let the mean value of ϕ over the lines of distance r from the origin be

$$\eta(r) = \frac{1}{2\pi}\int_0^{2\pi} \phi(r, \varphi)\, d\varphi$$

($\eta(r)$ is the mean value of F on the cylinder of radius r about l_0). $\eta(r)$ obviously depends only on the mean values $M(\rho)$ of w on the circles of radius ρ about the origin:

$$\eta(r) = 2\int_r^{\infty} \rho M(\rho)(\rho^2 - r^2)^{-\frac{1}{2}}\, d\rho.$$

Let $0 < z \leqq a$. Then

$$\begin{aligned}\frac{1}{2}\int_z^a \frac{\eta(r) - \eta(0)}{r^2}\, dr &= \int_z^a \frac{dr}{r^2}\int_r^{\infty} \rho M(\rho)(\rho^2 - r^2)^{-\frac{1}{2}}\, d\rho - \int_z^a \frac{dr}{r^2}\int_0^{\infty} M(\rho)\, d\rho \\ &= \int_z^{\infty} M(\rho)\left(\frac{(\rho^2 - z^2)^{\frac{1}{2}}}{\rho z} - \frac{1}{z}\right) d\rho - \frac{1}{z}\int_0^z M(\rho)\, d\rho \\ &\quad + \frac{1}{a}\int_a^{\infty} M(\rho)\left(1 - \frac{(\rho^2 - a^2)^{\frac{1}{2}}}{\rho}\right) d\rho + \frac{1}{a}\int_0^a M(\rho)\, d\rho.\end{aligned}$$

Now

$$\begin{aligned}&\left|\frac{1}{a}\int_a^{\infty} M(\rho)(1 - (1 - a^2\rho^{-2})^{\frac{1}{2}})\, d\rho + \frac{1}{a}\int_0^a M(\rho)\, d\rho\right| \\ &\qquad \leqq \left|\frac{1}{a}\int_a^{\infty} M(\rho)\, d\rho + \frac{1}{a}\int_0^{\infty} M(\rho)\, d\rho\right| \leqq \frac{2}{a}\int_0^{\infty} h(\rho)\, d\rho,\end{aligned}$$

$$\lim_{z\to 0} \frac{1}{z}\int_0^z M(\rho)\, d\rho = M(0) = w(0) = v(l_0),$$

and

$$\begin{aligned}&\int_z^{\infty} M(\rho)\left(\frac{(\rho^2 - z^2)^{\frac{1}{2}}}{\rho z} - \frac{1}{z}\right) d\rho \\ &\qquad = \int_1^{z^{-\frac{1}{2}}} M(\rho z)((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho + \int_{z^{-\frac{1}{2}}}^{\infty} M(\rho z)((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho,\end{aligned}$$

$$\left|\int_{z^{-\frac{1}{2}}}^{\infty} M(\rho z)((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho\right| \leqq -h(0)\int_{z^{-\frac{1}{2}}}^{\infty} ((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho \to 0$$

[7] Berichte Verh. Sächs. Akad. Wiss., vol. 69, pp. 262–277; discussions of this problem can also be found in a paper by Mader, Mathematische Zeitschrift, vol. 26, pp. 646–652, and in a paper by the author, Mathematische Annalen, vol. 109, p. 513 et seq.

for $z \to 0$. Also

$$\int_1^{z^{-\frac{1}{2}}} M(\rho z)((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho = M(\theta\sqrt{z}) \int_1^{z^{-\frac{1}{2}}} ((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho$$

$$\to M(0) \int_1^{\infty} ((1 - \rho^{-2})^{\frac{1}{2}} - 1)\, d\rho = w(0)(1 - \tfrac{1}{2}\pi)$$

for $z \to 0$. Consequently we obtain for $z \to 0$, $a \to \infty$ the required solution in the form

$$(13) \qquad v(l_0) = w(0) = -\frac{1}{\pi} \int_0^{\infty} \frac{\eta(r) - \eta(0)}{r^2}\, dr.$$

5. **Harmonic line functions determined by given data.** With the help of the last theorem we can easily find subsets S of straight lines with the property that the values of a harmonic line function v for the lines of S determine the values of v for a larger set uniquely.

Let S be a set of straight lines l in space with the property that with every l of S all parallels to l belong to S. Then S can be uniquely described by the set Σ of points of intersection of the lines of S with the plane at infinity π_∞. We write $S = S_\Sigma$.

DEFINITION. Let Σ be a set of points in the projective plane. Let $\bar{\Sigma}$ consist of all points P with the property that every line through P contains points of Σ. We shall call $\bar{\Sigma}$ the *linear extension* of Σ.

Examples. The linear extension of a straight line is the whole projective plane. The linear extension of a conic is its interior and boundary.

THEOREM 5.1. *Let the harmonic line function v be defined for all lines and be regular at infinity. Let v be given for all lines of a set $S = S_\Sigma$. Then v is uniquely determined for all lines of $S_{\bar{\Sigma}}$, where $\bar{\Sigma}$ is the linear extension of Σ.*

Proof. Let l be a line of $S_{\bar{\Sigma}}$ and P its point at infinity. Every line of π_∞ through P contains points of Σ, i.e., every plane parallel to l contains a family of parallel lines belonging to S, for which v is known. Thus the plane integral $F(\pi)$ of v is known for every plane parallel to l. From this $v(l)$ is uniquely determined according to Theorem 4.3 (and can actually be found with the help of Radon's formula (13)).

The notion of linear extension of a set Σ in a projective plane is of course closely connected with the notion of convex extension ("hull"). Let Σ be a closed set. A convex extension[8] of Σ is obtained by choosing a line not containing points of Σ as line at infinity and taking the intersection of all half-planes containing Σ in the remaining affine plane. It is easily proved that

(a) the linear extension $\bar{\Sigma}$ of Σ is the intersection of all convex extensions of Σ obtained by taking an arbitrary line exterior to Σ as line at infinity;

(b) if Σ is a connected set, all convex extensions of Σ are identical and equal to the linear extension $\bar{\Sigma}$;

(c) $\bar{\Sigma}$ contains the convex extensions of every connected subset of Σ.

[8] I use the term "convex extension" introduced by Dines (cf. Bull. Am. Math. Soc., vol. 42, p. 354) instead of "convex hull".

Special cases of Theorem 5.1 corresponding to the two examples of linear extensions given above are the following two statements.

v is uniquely determined for all straight lines, if given for all lines parallel to a given plane.

v is uniquely determined for all straight lines intersecting π_∞ in a point interior to a fixed conic ("time-like lines"), if given for all lines intersecting π_∞ in points of the conic ("characteristic lines").

One may interpret these results in terms of equation (2). A family of parallel lines in 3-space corresponds according to p. 307 to a characteristic 2-flat $x_4 - x_2 =$ const., $x_1 - x_3 =$ const. The set S of lines corresponds to a 3-dimensicnal manifold consisting of a family of 2-flats, which are parallel to the 2-flat $x_4 - x_2 = 0$, $x_1 - x_3 = 0$; Theorem 5.1 then states that a solution of (2), if given for the points of such a 3-dimensional manifold, is uniquely determined for a (in general) larger set of such 2-flats. Taking, for example, for S the set of lines parallel to the plane $x = 0$, we find that u is uniquely determined everywhere if given on the characteristic 3-flat $x_2 - x_4 = 0$, provided u has a certain regular behavior at infinity.

THEOREM 5.2. *Let $F(\pi)$ be a function of the planes π, defined for all π, which is such that*

(a) *F is bounded,*

(b) *F is continuous,*

(c) $\dfrac{\partial^2 F}{\partial p^2}$ *exists and is bounded for all π.*

(Here p denotes the distance of π from some fixed plane parallel to π.) Then the line function $v(l)$ defined by Radon's formula (12) *(for l_0 and correspondingly for all l) is harmonic. (It is not asserted that v is also regular at infinity. Thus F may possibly not be the plane integral of v.)*

Proof. For fixed r the line function $\eta(r) = \eta(r, l)$ is the mean value of the plane function $F(\pi)$ over the cylinders of radius r about the line l. Thus η is according to Theorem 3.4 for fixed r a harmonic function of l. If we can show that $\eta(r)$ and $r^{-2}[\eta(r) - \eta(0)]$ are uniformly bounded, $\int_0^\infty r^{-2}[\eta(r) - \eta(0)]dr$ will converge uniformly in l and therefore represent a harmonic line function as well. Now it is evident that every upper bound for F will be an upper bound for η. In order to prove that $\frac{1}{2}[\eta(r) - \eta(0)]$ is bounded, we determine the planes parallel to l by their distance p from l and the angle φ their normal includes with a fixed line perpendicular to l and define F for negative p by $F(p, \varphi) = F(-p, \varphi + \pi)$. (Note that our assumptions on F are invariant under rotation and translation.) Then

$$\eta(p) = \frac{1}{2\pi}\int_0^{2\pi} F(p, \varphi)\, d\varphi = \frac{1}{4\pi}\int_0^{2\pi} [F(p, \varphi) + F(-p, \varphi)]\, d\varphi,$$

$$\frac{\eta(p) - \eta(0)}{p^2} = \frac{1}{4\pi}\int_0^{2\pi} \frac{F(p, \varphi) - 2F(0, \varphi) + F(-p, \varphi)}{p^2}\, d\varphi;$$

the last expression is uniformly bounded since

$$\left| \frac{F(p, \varphi) - 2F(0, \varphi) + F(-p, \varphi)}{p^2} \right| \leqq \max \left| \frac{\partial^2 F}{\partial p^2} \right|.$$

Remark. It is obvious from the proof of Theorem 5.2 that assumption (c) can be replaced by the following weaker condition:

(c′) There exists a constant $A > 0$ such that for $|k| < A$

$$\frac{F(p + k) - 2F(p) + F(p - k)}{k^2}$$

is uniformly bounded for all planes π.

THEOREM 5.3. *Let Σ be a closed point set in π_∞. Let the line function $\bar{v}(l)$ be defined for the lines of the set S_Σ which intersect π_∞ in a point of Σ. Let $\bar{v}$ be continuous in S_Σ. Let there exist a monotonic function $h(r)$ such that $\int_0^\infty h(r)dr$ exists and such that*

(a) *$|\bar{v}(l)| < h(r)$ for a line l of S_Σ of distance r from the origin,*

(b) *$|\bar{v}(l_1) - 2\bar{v}(l) + \bar{v}(l_2)| \leqq h(r)p^2$, if l_1 and l_2 are the two parallels of l of distance p from l in any plane through l, where $|p| < A$.*

(c) *Let the integrals of $\bar{v}$ over any 2 families of parallels of S_Σ in the same plane be equal.*

Then there is a harmonic line function $v(l)$ defined for all l such that, for l in S_Σ, $v(l) = \bar{v}(l)$.

Proof. Let $F(\pi)$ denote the integral of $\bar{v}(l)$ over a family of parallel lines belonging to S_Σ in the plane π. According to assumptions (a) and (c) $F(\pi)$ is *uniquely* defined for the planes containing a point of Σ. We shall first prove that $F(\pi)$ is continuous. We say that a sequence of planes $\pi_1, \pi_2, \cdots$ converges towards a plane π, if the directions of the π_ν and their distance from the origin converge towards those of π. Let f_ν denote a family of parallels in π_ν belonging to S_Σ; let P_ν be their point at infinity. There is a subset of our sequence for which P_ν converges towards a limit point P. As $P_\nu \subset \Sigma$ and Σ is closed, $P \subset \Sigma$; also $P \subset \pi$, since $\lim_{\nu\to\infty} \pi_\nu = \pi$. Let $f \subset S_\Sigma$ denote the family of parallels through P in π. It is easily seen that one may establish a 1-1 correspondence between the lines l of f and the lines l_ν of f_ν such that $\lim_{\nu\to\infty} l_\nu = l$, and that a pair of parallels in f_ν corresponds to a pair of parallels in f with the same distance. As $F(\pi_\nu)$ is the integral of $\bar{v}(l_\nu)$ over the lines l_ν of f_ν and $F(\pi)$ the integral of $\bar{v}(l)$ over the lines l of f and as these improper integrals of $\bar{v}$ converge uniformly and $\bar{v}$ is continuous, it follows (by a similar argument as in the proof of Theorem 4.2) that $\lim_{\nu\to\infty} F(\pi_\nu) = F(\pi)$ for a suitable subsequence of the π_ν and therefore for every sequence converging towards π.

We now prove that $F(\pi)$ satisfies condition (c′) of Theorem 5.2. This condition concerns only the behavior of F for a family of parallel planes π, where p denotes, for example, the distance of π from the origin. Let P be a point of Σ

contained in π; let a line l through P be given by its distance x from a plane through the origin parallel to π and its distance y from a plane through P perpendicular to π; we have $\bar{v}(l) = \bar{v}(x, y)$. Then $F(p) = \int_{-\infty}^{+\infty} \bar{v}(p, y)dy$ and

$$
\begin{aligned}
| F(p + k) - 2F(p) + F(p - k) | \\
&= \left| \int_{-\infty}^{+\infty} (\bar{v}(p + k, y) - 2\bar{v}(p, y) + \bar{v}(p - k, y))\, dy \right| \\
&\leqq k^2 \int_{-\infty}^{+\infty} h((x^2 + y^2)^{\frac{1}{2}})\, dy \leqq 2k^2 \int_0^{\infty} h(r)\, dr.
\end{aligned}
$$

We thus have defined a plane function $F(\pi)$ satisfying conditions (a), (b), (c′) of Theorem 5.2 for all planes π containing a point of Σ. We can extend this definition to all planes. This extension is not uniquely determined; we shall give here one definite construction, which leads to the required results.

For this purpose we represent a plane π by its pole P with respect to the unit sphere. The points P corresponding to planes containing a point of Σ form a point set G, which consists of all points lying on a certain closed set of planes through the origin. $F(\pi)$ becomes a function $\phi(P)$ defined for the points of G. $\phi(P)$ will again be bounded and continuous (also at the origin, as $F(\pi)$ tends towards 0, if the distance of π from the origin tends towards infinity). G contains with every point P the whole line determined by P and the origin. If P, P', P'' are 3 points on the same line through the origin having distances p^{-1}, $(p + k)^{-1}$, $(p - k)^{-1}$, respectively, from the origin, then

$$\frac{1}{k^2} \,|\, \phi(P') - 2\phi(P) + \phi(P'') \,|$$

is uniformly bounded in G. If x, y, z denote rectangular coördinates of P, this expression may be written

$$
\begin{aligned}
\frac{1}{k^2} \Bigg(\phi \left(\frac{p}{p+k} x, \frac{p}{p+k} y, \frac{p}{p+k} z \right) - 2\phi(x, y, z) \\
+ \phi \left(\frac{p}{p-k} x, \frac{p}{p-k} y, \frac{p}{p-k} z \right) \Bigg).
\end{aligned}
\tag{14}
$$

Our task is to extend the definition of $\phi(P)$ to all points P, preserving the boundedness, continuity, and the boundedness of the expression (14). The complementary set Γ of G will consist of a certain number (finite or infinite) of connected subsets Γ_i. Every Γ_i is an open convex cone with vertex O, i.e., every 2 points P and Q in Γ_i can be joined by a straight line segment in Γ_i; for otherwise the segment PQ would contain a point of G and P and Q would then be separated by a whole plane belonging to G. We now distinguish 2 cases: (1) all points of Σ lie on a straight line, (2) Σ contains three non-collinear points.

We start with the second case. In that case G will consist of a number of planes through the origin, not all having one and the same line in common. Then there is a plane π_i through O having no point in common with Γ_i or its boundary except the origin. Every plane π parallel to π_i intersects Γ_i in a plane, open, finite, convex region γ, on the boundary of which ϕ is defined and continuous. We define $\phi(P)$ in the interior of γ as the potential function taking the prescribed values on the boundary. According to the maximum principle of potential functions[9] $\phi(P)$ will then be bounded in Γ_i and continuous in γ.

Let now P, P', P'' be three points in Γ_i on the same ray through O with distances p^{-1}, $(p + k)^{-1}$, $(p - k)^{-1}$, respectively, from O. Let us introduce a Cartesian coördinate system x, y, z with π_i as xy-plane. Let α be the distance of P from π_i. Then $z = \alpha$, $z = p(p + k)^{-1}\alpha$, $z = p(p - k)^{-1}\alpha$ are the three planes parallel to π_i through P, P', P'', respectively.

$$\phi(x, y, z), \quad \phi\left(\frac{p}{p+k}x, \frac{p}{p+k}y, \frac{p}{p+k}\alpha\right), \quad \phi\left(\frac{p}{p-k}x, \frac{p}{p-k}y, \frac{p}{p-k}\alpha\right)$$

are then 3 potential functions in x, y defined in the same region γ in which the plane $z = \alpha$ intersects Γ_i. Thus the expression (14) is a potential function as well and assumes the maximum of its absolute value on the boundary of γ, i.e., in a point of G. As this expression is bounded in G, the same bound applies in the interior of γ, i.e., in the point P.

In the remaining case that G consists of a number of planes all containing the same line through O, let Γ_i be a connected subset of Γ and π and π' the 2 planes bounding Γ_i. On every straight line segment in Γ_i which is perpendicular to the plane bisecting the angle between π and π' we define ϕ as a linear function taking the prescribed values in the points of intersection with π and π'. It is easily seen that our conditions on ϕ are again satisfied.

We now have a function F satisfying the assumptions (a), (b), (c′) of Theorem 5.2 defined for all planes. (The proof of the continuity of ϕ in Γ can be carried out along the same lines as that of the boundedness of (14).) If we define a line function $v(l)$ with the help of Radon's formula (13), $v(l)$ will be a harmonic line function defined for all l. If l belongs to S_2, then $v(l) = \bar{v}(l)$. For, for the planes parallel to l, $F(\pi)$ is *defined as the plane integral of* $\bar{v}$ over lines parallel to l and according to the proof of formula (13), this formula expresses the value of a line function at a particular line by its plane integrals over the planes parallel to the given line, provided the line function is continuous and regular at infinity.

Remark. Theorem 5.3 shows that a harmonic line function which is regular at infinity and satisfies the "Lipschitz condition" (b) can be continued over the whole space, provided (c) is also satisfied. Assumption (c) appears justified in view of Theorem 4.1, although its necessity is not proved, as the extended line function v may not be regular at infinity. A proof of the regularity of v at infinity would involve an estimation of Radon's expression (13), which is not

[9] The term "harmonic" function is not used in order to avoid confusion with harmonic line functions.

easy to perform.[10] The assumption (c) does not represent a restriction "im Kleinen" for $\bar{v}$, as long as the point set Σ has the property of being intersected by every line of π_∞ in only a finite number of points.

6. **Harmonic line functions as line integrals of a point function.** Let l_0 denote a line through the origin O. Let the lines l parallel to l_0 be determined by their distances ξ, η from two arbitrary fixed planes through l_0, which are perpendicular to one another. A line function $v(l)$ may then be written $v(\xi, \eta, l_0)$.

THEOREM 6. *Let the harmonic line function $v(l)$ be defined for all l. Let v have continuous first and second derivatives with respect to ξ and η. Let v, v_ξ, v_η be $O((\xi^2 + \eta^2)^{-\frac{1}{2}-\alpha})$ and $v_{\xi\xi}$, $v_{\xi\eta}$, $v_{\eta\eta}$ be $O((\xi^2 + \eta^2)^{-1-\alpha})$ uniformly in l_0, where α is a positive number. (These assumptions are independent of the coördinate system.) Then $v(l)$ may be represented as the integral of a point function $f(x, y, z)$ over the line l.*

Proof. Let $F(\pi)$ denote the integral of v over the plane π. We write $F(\pi) = F(p, \vartheta, \varphi)$, where p is the distance of π from the origin, ϑ the angle of the normal of π with the z-axis, and φ the angle of the line of intersection of π and the xy-plane with the y-axis. Let F be defined for negative p by

$$F(-p, \vartheta, \varphi) = F(p, \pi - \vartheta, \varphi + \pi). \tag{15}$$

We can easily prove that uniformly in ϑ and φ

$$\lim_{p\to\infty} p^{-1}F(p, \vartheta, \varphi) = 0, \tag{1}$$

$$F_p(p, \vartheta, \varphi) = O(p^{-\alpha}), \tag{2}$$

$$F_{p\vartheta}(p, \vartheta, \varphi) = O(p^{-\alpha}), \tag{3}$$

and that F_{pp} exists and is continuous.

(1) is an immediate consequence of the regular behavior of v at infinity (cf. Theorem 4.2). In order to prove (2) and (3) observe that in forming the derivatives F_p and $F_{p\vartheta}$ the angle φ is fixed and the planes π are parallel to a fixed line l_φ in the xy-plane. The plane integral $F(\pi)$ for the planes parallel to l_φ can be expressed with the help of the values of $v(l)$ for the lines parallel to l_φ alone. Let the lines parallel to l_φ be referred to a $\xi\eta$-coördinate system, where the η-axis is parallel to the z-axis. Then

$$F(p, \vartheta, \varphi) = \int_{-\infty}^{+\infty} v(p \sin \vartheta - s \cos \vartheta, p \cos \vartheta + s \sin \vartheta, l_\varphi)\, ds.$$

[10] In the next paragraph we shall encounter a similar difficulty. The proof of Theorem 6 would be very short if we could make use of the inverse of Radon's theorem, namely, that for a prescribed line function $\phi(p)$ in the plane, the point function w given by (12) has ϕ as line integral; but it is not even evident that the line integrals of w are convergent. Radon gives (loc. cit.) a set of conditions under which the inverse of his theorem holds, but they involve regularity conditions for the derivatives of the line function up to the third order, and the proof is complicated.

The derivatives of F with respect to p and ϑ can be immediately expressed as integrals of the derivatives of v with respect to ξ and η and (2) and (3) can be verified from our assumptions on v.

Let now $f(x, y, z)$ denote the mean value of $-(2\pi)^{-1}F_{pp}$ over the planes through the point (x, y, z). The function f is defined by F and therefore by v alone, since the value of F_{pp} is independent of the coördinate system.[11] We want to prove that $v(l)$ is the integral of f over the line l. As our assumptions are independent of the particular coördinate system, it is sufficient to prove this for the case that l is the z-axis. Using assumption (15), we have

$$f(0, 0, z) = -\frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi} F_{pp}(z\cos\vartheta, \vartheta, \varphi)\sin\vartheta\, d\vartheta\, d\varphi, \tag{16}$$

$$\begin{aligned}
\int_{-a}^{+a} f(0, 0, z)\, dz &= -\frac{1}{8\pi^2}\int_{-a}^{+a} dz\int_0^{2\pi} d\varphi\int_0^{\pi} d\vartheta\, F_{pp}(z\cos\vartheta, \vartheta, \varphi)\sin\vartheta \\
&= -\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_0^{\pi}\sin\vartheta\, d\vartheta\,\frac{F_p(a\cos\vartheta, \vartheta, \varphi) - F_p(-a\cos\vartheta, \vartheta, \varphi)}{\cos\vartheta} \\
&= -\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_0^{\pi}\sin\vartheta\, d\vartheta\,\frac{F_p(a\cos\vartheta, \frac{1}{2}\pi, \varphi) - F_p(-a\cos\vartheta, \frac{1}{2}\pi, \varphi)}{\cos\vartheta} \\
&\quad - \frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_0^{\pi}\sin\vartheta\, d\vartheta\,\frac{F_p(a\cos\vartheta, \vartheta, \varphi) - F_p(+a\cos\vartheta, \frac{1}{2}\pi, \varphi)}{\cos\vartheta} \\
&\quad + \frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_0^{\pi}\sin\vartheta\, d\vartheta\,\frac{F_p(-a\cos\vartheta, \vartheta, \varphi) - F_p(-a\cos\vartheta, \frac{1}{2}\pi, \varphi)}{\cos\vartheta} \\
&= A + B + C
\end{aligned}$$

(in easily understood notation). Here, using (3) and the theorem of mean value, we get

$$\begin{aligned}
B &= -\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_0^{\pi} d\vartheta\, F_{p\vartheta}(a\cos\vartheta, \theta, \varphi)\,\frac{\frac{1}{2}\pi - \vartheta}{\cos\vartheta}\sin\vartheta \\
&= O\left(\int_0^{2\pi} d\varphi\int_0^{\pi} d\vartheta\,\frac{1}{(a\cos\vartheta)^{\alpha}}\right) = O\left(\frac{1}{a^{\alpha}}\right),
\end{aligned}$$

where $\vartheta \leqq \theta \leqq \frac{1}{2}\pi$. Similarly it follows that $C = O(a^{-\alpha})$. By the substitution $\sigma = \cos\vartheta$ we obtain

$$\begin{aligned}
A &= -\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_{-1}^{+1}\frac{F_p(a\sigma, \frac{1}{2}\pi, \varphi) - F_p(-a\sigma, \frac{1}{2}\pi, \varphi)}{\sigma}\, d\sigma \\
&= -\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi\int_{-1}^{+1}\frac{F(a\sigma, \frac{1}{2}\pi, \varphi) + F(-a\sigma, \frac{1}{2}\pi, \varphi) - 2F(0, \frac{1}{2}\pi, \varphi)}{a\sigma^2}\, d\sigma \\
&\quad - \frac{1}{4\pi^2}\int_0^{2\pi}\frac{F(a, \frac{1}{2}\pi, \varphi) + F(-a, \frac{1}{2}\pi, \varphi) - 2F(0, \frac{1}{2}\pi, \varphi)}{a}\, d\varphi.
\end{aligned}$$

[11] The expression for f is chosen in accordance with Radon's solution of the problem of determining a point function f in space by its plane integrals. Cf. Radon, loc. cit.

If $\eta(\sigma)$ is defined by $\eta(\sigma) = (2\pi)^{-1} \int_0^{2\pi} F(\sigma, \frac{1}{2}\pi, \varphi)d\varphi$, we have $\eta(\sigma) = \eta(-\sigma)$, $\lim_{\sigma\to\infty} \eta(\sigma)\sigma^{-1} = 0$ and

$$\begin{aligned} A &= -\frac{1}{4\pi}\int_{-1}^{+1} \frac{\eta(a\sigma) + \eta(-a\sigma) - 2\eta(0)}{a\sigma^2}\, d\sigma - \frac{1}{\pi}\frac{\eta(a) - \eta(0)}{a} \\ &= -\frac{1}{\pi}\int_0^a \frac{\eta(\sigma) - \eta(0)}{\sigma^2}\, d\sigma - \frac{1}{\pi}\frac{\eta(a) - \eta(0)}{a} \end{aligned}$$

and correspondingly

$$\int_{-\infty}^{+\infty} f(0, 0, z)\, dz = \lim_{a\to\infty} A = -\frac{1}{\pi}\int_0^\infty \frac{\eta(\sigma) - \eta(0)}{\sigma^2}\, d\sigma.$$

As $\eta(\sigma)$ is the mean value of F over the cylinder of radius σ about the z-axis, the second term in the last equation represents according to (13) the value of v for the z-axis. Thus our proof is finished.

We conclude with the remark that the function $f(x, y, z)$ constructed here satisfies the assumptions of Theorem 3.3. Indeed we may derive from (16) and (2) and (3) by "partial integration"[12]

$$\begin{aligned} f(0, 0, z) = &-\frac{1}{8\pi^2}\int_0^{2\pi} \frac{F_p(-z, \pi, \varphi) - F_p(z, 0, \varphi)}{z}\, d\varphi \\ &-\frac{1}{8\pi^2}\int_0^{2\pi} d\varphi \int_0^\pi \frac{F_{p\vartheta}(z\cos\vartheta, \vartheta, \varphi)}{z}\, d\vartheta = O(z^{-1-\alpha}). \end{aligned}$$

UNIVERSITY OF KENTUCKY.

[12] In the sense in which the expression is used by M. Brendel, Mathematische Annalen, vol. 55, p. 248, and N. J. Hatzidakis, ibid., vol. 57, p. 134.

A COTANGENT ANALOGUE OF CONTINUED FRACTIONS

By D. H. Lehmer

The continued iteration of a rational function $f(x, y)$ of two variables provides an algorithm for the expression of a real number as a sequence of rational numbers. Thus the function

$$f(x_1, f(x_2, f(x_3, \cdots))) \tag{1}$$

becomes an infinite series for $f(x, y) = x + y$ and an infinite product for $f(x, y) = xy$. For $f(x, y) = x + 1/y$ we obtain the regular continued fraction

$$x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cdots}} = x_1 + \frac{1|}{|x_2} + \frac{1|}{|x_3} + \cdots.$$

By far the most frequently used function is $f(x, y) = x + y/c$, which gives the "power series"

$$x_1 + \cfrac{x_2 + \cfrac{x_3 + \cdots}{c}}{c} = x_1 + \frac{x_2}{c} + \frac{x_3}{c^2} + \cdots,$$

where the x's are the coefficients, used when $c = 10$ for the decimal representation of real numbers.[1] The algorithm associated with $f(x, y) = x(1 - y)$ has been discussed by T. A. Pierce.[2]

This paper is concerned with the case of

$$f(x, y) = (xy + 1)/(y - x) = \cot(\operatorname{arc\,cot} x - \operatorname{arc\,cot} y),$$

so that (1) becomes the function

$$\cot(\operatorname{arc\,cot} x_1 - \operatorname{arc\,cot} x_2 + \operatorname{arc\,cot} x_3 - \cdots).$$

This function, despite its aspect, is no more transcendental than a regular continued fraction and both functions have many properties in common. Furthermore, in order to obtain sequences of rational approximations to a real number, we specialize the x's to be integers, as in the continued fraction, and consider therefore expressions of the form

$$\cot \sum_{\nu=0} (-1)^\nu \operatorname{arc\,cot} n_\nu, \tag{2}$$

Received November 24, 1937.

[1] This use of the function $x + y/c$ is at least 4000 years old. See Amer. Jour. of Semitic Languages and Literature, vol. 36(1920), No. 4. The Babylonians used $c = 60$.

[2] Amer. Math. Monthly, vol. 36(1929), pp. 523–525.

where the n_ν are integers. This expression will be called a "continued cotangent", and we shall use the adjective "finite" or "infinite" according as the series in (2) terminates or not. Although finite and infinite sums of arc cotangents of integers have been considered many times, no systematic treatment of such sums appears to have been given.

Definition of a regular continued cotangent. The continued cotangent (2) will be said to be regular if

(a) n_ν is an integer[3] $\geqq 0$ for $\nu \geqq 0$.

(b) If (2) is finite and if n_k is the last n, then

$$n_k > n_{k-1}^2 + n_{k-1} + 1. \tag{3}$$

In all other cases

$$n_\nu \geqq n_{\nu-1}^2 + n_{\nu-1} + 1. \tag{4}$$

The principal value of arc cotangent n_ν is understood. In fact, since n_ν is non-negative,

$$0 < \text{arc cot } n_\nu \leqq \tfrac{1}{2}\pi.$$

The inequalities (3) and (4) seem at first sight unnatural. They are, however, the analogues of the inequalities

$$q_k > 1, \tag{3'}$$

$$q_\nu \geqq 1 \tag{4'}$$

for the incomplete quotients of the continued fraction

$$q_0 + \frac{1|}{|q_1} + \frac{1|}{|q_2} + \cdots,$$

which terminates with $\cdots + \dfrac{1|}{|q_k}$ or is infinite. The reason for insisting on the stronger inequality (3) in the case of a finite continued cotangent is the same as the reason for (3′) in the continued fraction: to insure for every rational number a unique expansion. As a matter of fact, if (4) held for n_k but not (3), so that

$$n_k = n_{k-1}^2 + n_{k-1} + 1, \tag{5}$$

then the last two terms of (2) could be replaced by a single term, since

$$\text{arc cot } n_{k-1} - \text{arc cot } (n_{k-1}^2 + n_{k-1} + 1) = \text{arc cot } (n_{k-1} + 1),$$

just as in continued fractions we write

$$\frac{1|}{|q_{k-1}} + \frac{1|}{|1} = \frac{1|}{|q_{k-1} + 1}.$$

[3] As in continued fractions we might allow n_0 to be negative. However, this extra generality is non-essential for our purposes.

Hence (3) may as well be assumed. It is perhaps worth noting that this contraction of the last two terms cannot be repeated in the continued cotangent any more than in continued fractions. In fact we would need to have as a counterpart of (5)

$$n_{k-1} + 1 = n_{k-2}^2 + n_{k-2} + 1.$$

This violates (4).

THEOREM 1. *Every infinite regular continued cotangent converges.*

Proof. We need merely to note that

$$\text{arc cot } n_0 - \text{arc cot } n_1 + \text{arc cot } n_2 - \cdots \tag{6}$$

form an alternating series of terms monotonically decreasing in absolute value in view of (4). Since (6) converges to a positive quantity, the cotangent of (6) exists, and this proves the theorem. In fact, it is easy to see that (6) converges not only absolutely but with tremendous rapidity, more rapidly, indeed, than the series

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{16} + \frac{1}{256} + \frac{1}{65536} + \cdots + \frac{1}{2^{2^n}} + \cdots,$$

in view of (4) and the inequality

$$\text{arc cot } n_\nu < \frac{1}{n_\nu}.$$

This rapidity of convergence is a feature of the continued cotangent not enjoyed by the continued fraction. The least rapidly converging continued fraction may be said to be

$$\frac{\sqrt{5}-1}{2} = 0 + \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|1} + \cdots, \tag{7}$$

whereas the least rapidly converging continued cotangent is

$$\begin{aligned} \xi = \cot (&\text{arc cot } 0 - \text{arc cot } 1 + \text{arc cot } 3 - \text{arc cot } 13 + \text{arc cot } 183 \\ &- \text{arc cot } 33673 + \text{arc cot } 1133904603 - \cdots), \end{aligned} \tag{8}$$

in which

$$n_{\nu+1} = n_\nu^2 + n_\nu + 1.$$

Uniqueness theorem. Theorem 1 guarantees that every continued cotangent represents a real positive number. Before treating the inverse problem of finding the continued cotangent expansion of a given number, we prove the following uniqueness theorem.

THEOREM 2. *Two regular continued cotangents can be equal only if they are identically equal.*

Proof. Let

$$\cot \sum_{\nu=0} (-1)^\nu \operatorname{arc\,cot} n_\nu = \cot \sum_{\nu=0} (-1)^\nu \operatorname{arc\,cot} m_\nu \tag{9}$$

be two equal regular continued cotangents, and suppose, if possible, that $n_\nu = m_\nu$ does not hold for all ν. Then there exists a first instance, $\nu = r$, where $n_r \neq m_r$ while $n_\nu = m_\nu$ for $\nu < r$ if $r \neq 0$. Then from (9) we have

$$\sum_{\lambda=0} (-1)^\lambda \operatorname{arc\,cot} n_{r+\lambda} = \sum_{\lambda=0} (-1)^\lambda \operatorname{arc\,cot} m_{r+\lambda} = S. \tag{10}$$

Since $n_r \neq m_r$, at least one of these sums contains two or more terms. Let this sum be the left one, so that

$$\begin{aligned} S = \sum_{\lambda=0} (-1)^\lambda \operatorname{arc\,cot} n_{r+\lambda} &\geqq \operatorname{arc\,cot} n_r - \operatorname{arc\,cot} n_{r+1} \\ &= \operatorname{arc\,cot}\left(n_r + \frac{n_r^2+1}{n_{r+1}-n_r}\right) \geqq \operatorname{arc\,cot}(n_r+1). \end{aligned} \tag{11}$$

In fact, the first $\geqq$ sign reads $=$ only if arc cot n_{r+1} is the last term of the left member of (10). In this case, however, (3) applies, so that

$$n_{r+1} > n_r^2 + n_r + 1, \quad \text{or} \quad \frac{n_r^2+1}{n_{r+1}-n_r} < 1.$$

Therefore the second $\geqq$ sign in (11) reads $>$ in case the first reads $=$. That is,

$$S > \operatorname{arc\,cot}(n_r + 1). \tag{12}$$

But since the left member of (10) contains at least two terms,

$$S < \operatorname{arc\,cot} n_r. \tag{13}$$

We may now show that the right member of (10) contains two or more terms; otherwise we could write from (10), (12) and (13)

$$\operatorname{arc\,cot}(n_r + 1) < S = \operatorname{arc\,cot} m_r < \operatorname{arc\,cot} n_r.$$

That is, $n_r + 1 > m_r > n_r$. But this is impossible, since these letters are integers.[4] We conclude, therefore, that both members of (10) contain two or more terms. Hence not only is $S <$ arc cot m_r, so that

$$m_r < n_r + 1, \tag{14}$$

but also, since the reasoning used to establish (12) may be now applied to the m's, $S >$ arc cot $(m_r + 1)$. Combining this with (13), we have $n_r < m_r + 1$. Finally in view of (14) we may write $n_r - 1 < m_r < n_r + 1$. But this contradicts $m_r \neq n_r$. Hence the theorem is proved.

[4] This is the first place that this part of the definition of the regular continued cotangent is used.

Arc cotangent algorithm. We now describe an algorithm, analogous to that of Euclid, for generating from a given real positive number its regular continued cotangent expansion.

Let x be the given positive number. We define two sets of numbers x_ν and n_ν ($\nu = 0, 1, 2, \cdots$) called respectively the ν-th complete and incomplete cotangent of x as follows.[5]

$$\begin{aligned} x_0 &= x, & n_0 &= [x_0], \\ x_1 &= \frac{x_0 n_0 + 1}{x_0 - n_0}, & n_1 &= [x_1], \\ x_2 &= \frac{x_1 n_1 + 1}{x_1 - n_1}, & n_2 &= [x_2], \\ &\cdots\cdots\cdots & & \\ x_{\nu+1} &= \frac{x_\nu n_\nu + 1}{x_\nu - n_\nu}, & n_{\nu+1} &= [x_{\nu+1}]. \end{aligned} \tag{15}$$

This algorithm is to be continued as long as $x_{\nu+1}$ exists, that is, as long as x_ν is not an integer $n_\nu = [x_\nu]$. We next prove

THEOREM 3. *The continued cotangent*

$$\cot \sum_{\nu=0} (-1)^\nu \operatorname{arc\,cot} n_\nu , \tag{16}$$

where the sum extends over all the incomplete cotangents n_ν of x, is regular.

Proof. Obviously (a) is satisfied. To show that (b) is satisfied we set $x_\nu = n_\nu + \epsilon_\nu$, where $0 < \epsilon_\nu < 1$. Then (15) becomes

$$x_{\nu+1} = \frac{n_\nu^2 + 1}{\epsilon_\nu} + n_\nu > n_\nu^2 + n_\nu + 1. \tag{17}$$

Hence

$$[x_{\nu+1}] = n_{\nu+1} \geqq n_\nu^2 + n_\nu + 1,$$

so that (4) is satisfied for $\nu \neq k - 1$. For $\nu = k - 1$ we have from (17)

$$x_k = n_k > n_{k-1}^2 + n_{k-1} + 1,$$

which is (3). Hence the theorem is true.

THEOREM 4. *If $n_0, n_1, n_2, \cdots$ are generated by x, then*

$$\sum_{\nu=0}^{\mu-1} (-1)^\nu \operatorname{arc\,cot} n_\nu = \operatorname{arc\,cot} x - (-1)^\mu \operatorname{arc\,cot} x_\mu . \tag{18}$$

Remark. This theorem justifies the name "complete cotangent" for x_μ.

Proof. Since

$$x_{\nu+1} = \frac{x_\nu n_\nu + 1}{x_\nu - n_\nu},$$

5 Here, as usual, $[x]$ means the greatest integer $\leqq x$.

we have

$$(-1)^\nu \text{ arc cot } n_\nu = (-1)^\nu (\text{arc cot } x_{\nu+1} + \text{arc cot } x_\nu).$$

Setting $\nu = 0, 1, 2, \cdots, \mu - 1$ and adding, we get the theorem.

THEOREM 5.

$$x = \cot \sum_{\nu=0} (-1)^\nu \text{ arc cot } n_\nu, \tag{19}$$

where the sum extends over all incomplete cotangents n_ν generated by x.

Proof. In case there exists only a finite number of n's, the last being n_k, we may set $\mu = k$ in (18) and transpose the term $(-1)^k$ arc cot x_k. Taking the cotangent of both sides we obtain (19).

In case an infinite number of n's are generated by x we can write in view of (18) and (4),

$$\lim_{\mu=\infty} \sum_{\nu=0}^{\mu-1} (-1)^\nu \text{ arc cot } n_\nu = \text{arc cot } x - \lim_{\mu=\infty} (-1)^\mu \text{ arc cot } x_\mu = \text{arc cot } x.$$

Hence in this case also

$$x = \sum_{\nu=0}^{\infty} (-1)^\nu \text{ arc cot } n_\nu.$$

THEOREM 6. *Every positive number has a unique regular continued cotangent expansion.*

Proof. The existence of such an expansion follows from the arc cotangent algorithm and Theorem 5, while the uniqueness is provided by Theorem 2.

THEOREM 7. *The number x is rational or irrational according as its continued cotangent expansion* (19) *is finite or not.*

Proof. If (19) is finite, it follows from the addition theorem of the cotangent function that x is rational. This may be seen otherwise. In fact, if x were irrational, so also would be $x_1, x_2, \cdots$. Hence there could not exist a k for which x_k is an integer to terminate the algorithm.

If (19) is infinite, then x is irrational. In fact, suppose that $x = p/q$, where p and q are integers. It follows that $x_\nu = p_\nu/q_\nu$ is also rational for every ν. From (15)

$$x_{\nu+1} = \frac{p_{\nu+1}}{q_{\nu+1}} = \frac{p_\nu n_\nu + q_\nu}{p_\nu - n_\nu q_\nu} = \frac{p_\nu n_\nu + q_\nu}{r_\nu},$$

where, since $n_\nu = [x_\nu] = [p_\nu/q_\nu]$, the denominator r_ν is the remainder on division of p_ν by q_ν, so that $r_\nu < q_\nu$. Since we may suppose that the fraction $p_{\nu+1}/q_{\nu+1}$ is in its lowest terms, we have the inequality

$$q_{\nu+1} \leqq r_\nu < q_\nu$$

for every ν. But this implies the existence of an infinite sequence $q_1, q_2, \cdots$ of strictly decreasing positive integers, and this is absurd. Hence x is irrational.

If x is a rational number p/q, the successive numerators p_ν and the denominators q_ν of x_ν can be found as in the greatest common divisor process as follows:

$$\begin{array}{lll}
p\ \ = n_0q + q_1 & (0 \leqq q_1 < q), & pn_0 + q = p_1\,, \\
p_1 = n_1q_1 + q_2 & (0 \leqq q_2 < q_1), & p_1n_1 + q_1 = p_2\,, \\
p_2 = n_2q_2 + q_3 & (0 \leqq q_3 < q_2), & p_2n_2 + q_2 = p_3\,, \\
\cdots\cdots\cdots & \cdots\cdots\cdots & \cdots\cdots\cdots \\
p_\nu = n_\nu q_\nu + q_{\nu+1} & (0 \leqq q_{\nu+1} < q_\nu), & p_\nu n_\nu + q_\nu = p_{\nu+1}\,, \\
\cdots\cdots\cdots & & \\
p_k = n_kq_k\,. & &
\end{array}$$

In general, p_ν will not be prime to q_ν. In fact, any factor which they may have in common will be a common factor of $p_{\nu+1}$ and $q_{\nu+1}$ and hence of all further p's and q's. For example, for $x = 65/37$, we find the following values of p_ν, q_ν, n_ν, and the greatest common divisor δ_ν of p_ν and q_ν.

ν	0	1	2	3
p_ν	65	102	334	6030
q_ν	37	28	18	10
n_ν	1	3	18	603
δ_ν	1	2	2	10

Hence $65/37 = \cot(\operatorname{arc\,cot} 1 - \operatorname{arc\,cot} 3 + \operatorname{arc\,cot} 18 - \operatorname{arc\,cot} 603)$.

Convergents. Let $n_0, n_1, n_2, \cdots$ be the incomplete cotangents generated by x. Then the curtate expansion of μ terms

$$\sigma_\mu(x) = \cot \sum_{\nu=0}^{\mu-1} (-1)^\nu \operatorname{arc\,cot} n_\nu$$

is called the μ-th convergent of x. It is clearly a rational number depending only on μ and x. The following expression relates x, $\sigma_\mu(x)$, and the complete cotangent x_μ by (18):

$$\sigma_\mu(x) = \cot(\operatorname{arc\,cot} x - (-1)^\mu \operatorname{arc\,cot} x_\mu) = \frac{(-1)^\mu x_\mu x + 1}{(-1)^\mu x_\mu - x}. \tag{20}$$

THEOREM 8. *If the integers A_ν and B_ν are defined by*

$$\begin{aligned} A_0 &= 1, \quad & A_{\nu+1} &= A_\nu n_\nu - (-1)^\nu B_\nu\,, \\ B_0 &= 0, \quad & B_{\nu+1} &= B_\nu n_\nu + (-1)^\nu A_\nu\,, \end{aligned} \tag{21}$$

then the μ-th complete cotangent is given by

$$x_\mu = (-1)^\mu \frac{A_\mu x + B_\mu}{A_\mu - B_\mu x}, \tag{22}$$

and the μ-th convergent $\sigma_\mu(x)$ is given by

$$\sigma_\mu(x) = A_\mu/B_\mu\,. \tag{23}$$

Proof. Formula (22) is easily established by induction. In fact (22) holds for $\mu = 0$, since $A_0 = 1, B_0 = 0, x_0 = x$. If it is true for $\mu = \nu$, we may write by (15) and (21)

$$x_{\nu+1} = \frac{(-1)^\nu(A_\nu x + B_\nu)n_\nu + A_\nu - B_\nu x}{(-1)^\nu(A_\nu x + B_\nu) - n_\nu(A_\nu - B_\nu x)}$$
$$= (-1)^{\nu+1}\frac{(A_\nu n_\nu - (-1)^\nu B_\nu)x + B_\nu n_\nu + (-1)^\nu A_\nu}{A_\nu n_\nu - (-1)^\nu B_\nu - (B_\nu n_\nu + (-1)^\nu A_\nu)x}$$
$$= (-1)^{\nu+1}\frac{A_{\nu+1}x + B_{\nu+1}}{A_{\nu+1} - B_{\nu+1}x},$$

so that the induction is complete. Having established (22), we see that (23) follows from (20). In fact,

$$\sigma_\mu(x) = \left\{\frac{A_\mu x^2 + B_\mu x}{A_\mu - B_\mu x} + 1\right\}\Big/\left\{\frac{A_\mu x + B_\mu}{A_\mu - B_\mu x} - x\right\} = A_\mu/B_\mu .$$

The numbers A_μ and B_μ are, of course, the analogues of the numerator and denominator of the μ-th convergent of the regular continued fraction. However, the recurrence formulas (21) are of a different nature, A_μ or B_μ depending not on the preceding A's or B's, but on the preceding A and B. This fact allows one to give an explicit formula for A_μ and B_μ in terms of the first μ incomplete cotangents $n_0, n_1, \cdots, n_{\mu-1}$.

$$A_0 = 1, \qquad B_0 = 0,$$
$$A_1 = n_0, \qquad B_1 = 1,$$
$$A_2 = n_0n_1 + 1, \qquad B_2 = n_1 - n_0,$$
$$A_3 = n_0n_1n_2 + n_0 - n_1 + n_2, \qquad B_3 = n_0n_1 - n_0n_2 + n_1n_2 + 1,$$
$$A_4 = n_0n_1n_2n_3 + n_0n_1 + n_1n_2 + n_1n_3 + n_2n_3 - n_1n_3 - n_0n_2 + 1,$$
$$B_4 = n_0n_1n_3 - n_0n_2n_3 + n_1n_2n_3 - n_0n_1n_2 + n_1 + n_3 - n_0 - n_2 .$$

The general formula for the A_μ and the B_μ is given by

THEOREM 9.

$$A_\mu + iB_\mu = (n_0 + i)(n_1 - i)(n_2 + i)(n_3 - i)\cdots(n_{\mu-1} + (-1)^{\mu-1}i)$$
$$(24) \qquad = \prod_{\nu=0}^{\mu-1}(n_\nu + (-1)^\nu i) \qquad (i^2 = -1).$$

In other words, if S_ν denotes the sum of the products of $(-1)^t n_t$ taken ν at a time, then for $\mu > 0$

$$(-1)^{[\frac{1}{2}\mu]}A_\mu = S_\mu - S_{\mu-2} + S_{\mu-4} - \cdots,$$
$$(-1)^{[\frac{1}{2}\mu]}B_\mu = S_{\mu-1} - S_{\mu-3} + S_{\mu-5} - \cdots.$$

Proof. Formula (24) is easily established by induction, if we use (21). It also follows readily from

$$\operatorname{arc\,cot} u = \frac{1}{2i} \log \frac{u+i}{u-i}.$$

THEOREM 10.

$$(25)\qquad \begin{vmatrix} A_\mu & B_\mu \\ A_{\mu+1} & B_{\mu+1} \end{vmatrix} = (-1)^\mu (A_\mu^2 + B_\mu^2) = (-1)^\mu \prod_{\nu=0}^{\mu-1} (n_\nu^2 + 1).$$

Proof. The first equality follows at once from (21) while the second equality is obtained by taking the squares of the absolute values of both sides of (24).

THEOREM 11.

$$(26)\quad A_\mu A_{\mu+1} + B_\mu B_{\mu+1} = \begin{vmatrix} A_\mu & iB_\mu \\ iB_{\mu+1} & A_{\mu+1} \end{vmatrix} = n_\mu (A_\mu^2 + B_\mu^2) = n_\mu \prod_{\nu=0}^{\mu-1} (n_\nu^2 + 1).$$

Proof. The theorem follows at once from (21) and (25).

For example, the values of A_ν, B_ν for $x = 65/37$ are given in the following table.

ν	0	1	2	3	4
n_ν	1	3	18	603	
A_ν	1	1	4	70	42250
B_ν	0	1	2	40	24050

Here we find that $A_4/B_4 = 65/37$ and that A_4 and B_4 have the common factor $650 = (n_0^2 + 1)(n_2^2 + 1)$.

As a second example, we give the elements for $x = 6954069/2559142$.

ν	0	1	2	3	4
p_ν	6954069	16467280	133574025	9886258850	
q_ν	2559142	1835785	1781000	1780025	
n_ν	2	8	74	5554	
A_ν	1	2	17	1252	6954069
B_ν	0	1	6	461	2559142

In this example A_ν and B_ν have no common factor. It is clear from (25) that any factor common to A_ν and B_ν will divide $(n_0^2 + 1)(n_1^2 + 1) \cdots (n_{\nu-1}^2 + 1)$, and this factor will also be common to $(A_{\nu+1}, B_{\nu+1})$ by (21) and hence to all the further pairs (A, B).

THEOREM 12. *The convergents $\sigma_\nu(x)$ approach x with errors which are alternately positive and negative, but whose absolute values tend steadily to zero and are less than*

$$(x\sigma_\nu + 1) \tan \varphi_\nu,$$

where φ_ν is the smaller of $[x]^{-2^\nu}$ and $3^{-2^\nu-2}$.

Proof. By definition of σ_ν,

$$\text{(27)} \qquad \operatorname{arc\,cot} x = \operatorname{arc\,cot} \sigma_\nu + (-1)^\nu \{\operatorname{arc\,cot} n_\nu - \operatorname{arc\,cot} n_{\nu+1} + \cdots\}.$$

Since arc cot u is a decreasing function of u, we have

$$\text{(28)} \qquad (-1)^{\nu+1}(x - \sigma_\nu) \geqq 0,$$

which implies the first statement of the theorem. Moreover, by (27) and (4),

$$|\operatorname{arc\,cot} x - \operatorname{arc\,cot} \sigma_\nu| \leqq \operatorname{arc\,cot} n_\nu < n_\nu^{-1} < n_{\nu-1}^{-2} < \cdots < n_2^{-2^{\nu-2}} < n_0^{-2^\nu}.$$

Hence if $n_0 = [x] > 1$, we may write

$$|\operatorname{arc\,cot} x - \operatorname{arc\,cot} \sigma_\nu| < [x]^{-2^\nu}.$$

If $n_0 = [x] \leqq 1$, then, by (4), $n_1 \geqq 1$, $n_2 \geqq 3$. Therefore in this case

$$|\operatorname{arc\,cot} x - \operatorname{arc\,cot} \sigma_\nu| < 3^{-2^{\nu-2}}.$$

Hence in either case

$$|\operatorname{arc\,cot} x - \operatorname{arc\,cot} \sigma_\nu| < \varphi_\nu,$$

and the final statement of the theorem follows by taking the tangent of both sides of this inequality. It remains to show that the absolute value of the error tends steadily to zero. Denoting this absolute value by Δ_ν, we have by (28) and (20)

$$\text{(29)} \qquad \Delta_\nu = |x - \sigma_\nu| = (-1)^{\nu+1}(x - \sigma_\nu) = \frac{x^2 + 1}{x_\nu - (-1)^\nu x}.$$

To show that Δ_ν is greater than $\Delta_{\nu+1}$, it suffices to show that

$$\text{(30)} \qquad (1 + x^2)(\Delta_{\nu+1}^{-1} - \Delta_\nu^{-1}) = x_{\nu+1} - x_\nu + (-1)^\nu 2x$$

is positive. From (15) and (4)

$$x_{\nu+1} > n_\nu x_\nu + 1, \qquad n_\nu \geqq 3 \qquad (\nu \geqq 2).$$

Hence $x_{\nu+1} - x_\nu > x_\nu - x_{\nu-1}$. It follows from (30) that

$$(1 + x^2)(\Delta_{\nu+1}^{-1} - \Delta_\nu^{-1}) > x_2 - x_1 - 2x.$$

To show that the right member is positive we separate two cases. If $x > 1$, then $n_1 \geqq 3$, $x_2 > 3x_1 + 1$, $x_1 > n_0 x + 1 \geqq x + 1$. Hence in this case

$$x_2 - x_1 - 2x > 2(x + 1) + 1 - 2x = 3 > 0.$$

If $x < 1$, let $x^{-1} = \delta > 1$. Then $x_1 = \delta$, $n_1 = \delta - \epsilon$ $(0 < \epsilon < 1)$, $x_2 = \epsilon^{-1}[\delta(\delta - \epsilon) + 1]$. Therefore

$$x_2 - x_1 - 2x = (\delta^2 + 1)(\delta - 2\epsilon)/\delta\epsilon.$$

If $\delta > 2$, this is positive. If $1 < \delta < 2$ so that $n_1 = \delta - \epsilon = 1$, we have

$$x_2 - x_1 - 2x = (\delta^2 + 1)(1 - \epsilon)/\delta\epsilon > 0.$$

This completes the proof of the theorem.

The expression of a regular continued cotangent as an irregular continued fraction. The partial cotangents n_ν of a number x may be used to represent x by an irregular continued fraction of special type as the following theorem shows.

THEOREM 13. *If $n_0, n_1, \cdots$ are the partial cotangents generated by a real positive number x, then*

$$x = n_0 + \frac{n_0^2 + 1|}{|n_1 - n_0} + \frac{n_1^2 + 1|}{|n_2 - n_1} + \frac{n_2^2 + 1|}{|n_3 - n_2} + \cdots. \tag{31}$$

Proof. Let ϵ_ν be the fractional part of x_ν so that $x_\nu = n_\nu + \epsilon_\nu$. Substituting for $x_{\nu+1}$ and x_ν in (15) and solving for ϵ_ν, we obtain

$$\epsilon_\nu = \frac{n_\nu^2 + 1}{n_{\nu+1} - n_\nu + \epsilon_{\nu+1}}.$$

Setting $\nu = 0, 1, 2, \cdots$ in succession, we see that (31) follows from $x = n_0 + \epsilon_0$. It is clear also that the numbers $A_{\mu+1}$ and $B_{\mu+1}$ are the numerator and denominator of the μ-th convergent of (31).

Regular continued fraction for ξ. The number ξ defined by (8) may be expressed as a regular continued fraction as follows. Let $n_0, n_1, n_2, \cdots$ be the partial cotangents of ξ so that

$$n_{\nu+1} - n_\nu = n_\nu^2 + 1. \tag{32}$$

We define integers a_ν by

$$a_0 = 1, \quad a_1 = 1, \quad a_2 = 2, \quad a_3 = 5, \quad a_4 = 34, \quad a_5 = 985$$

and in general

$$a_{\nu+1} = (n_\nu + n_{\nu-1} + 1)a_{\nu-1} \qquad (\nu \geqq 1), \tag{33}$$

so that

$$a_{\nu+1} = (n_\nu + n_{\nu-1} + 1)(n_{\nu-2} + n_{\nu-3} + 1)(n_{\nu-4} + n_{\nu-5} + 1) \cdots, \tag{34}$$

where the last factor is $n_1 + n_0 + 1 = 2$ or $n_2 + n_1 + 1 = 5$ according as ν is even or odd. Then it is true that

$$a_{\nu+1}a_\nu = n_{\nu+1} - n_\nu = n_\nu^2 + 1. \tag{35}$$

This fact is true for $\nu = 0$, since $n_0^2 + 1 = 1$, while $a_0a_1 = 1$. If it is true for $\nu = k - 1$, it may be shown true for $\nu = k$ as follows:

$$\begin{aligned} a_{k+1}a_k = \frac{a_{k+1}}{a_{k-1}} a_k a_{k-1} &= (n_k + n_{k-1} + 1)(n_k - n_{k-1}) \\ &= n_k^2 + n_k - n_{k-1}^2 - n_{k-1} = n_k^2 + 1. \end{aligned}$$

This establishes (35). Returning to (31) and using (32) and (35), we obtain

$$\begin{aligned} \xi &= 0 + \frac{a_0 a_1|}{|a_0 a_1} + \frac{a_1 a_2|}{|a_1 a_2} + \frac{a_2 a_3|}{|a_2 a_3} + \cdots \\ &= \frac{1|}{|a_0} + \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots \\ &= \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|2} + \frac{1|}{|5} + \frac{1|}{|34} + \frac{1|}{|985} \\ &\qquad + \frac{1|}{|1151138} + \frac{1|}{|1116929202845} + \cdots. \end{aligned} \tag{36}$$

The successive convergents C_ν/D_ν to ξ are

$$\frac{1}{1}, \frac{1}{2}, \frac{3}{5}, \frac{16}{27}, \frac{547}{923}, \frac{538811}{909182}, \frac{620245817465}{1046593950039}, \cdots.$$

In decimals we have

$$\xi = \begin{matrix} .59263 & 27182 & 01636 & 19710 & 40786 & 04995 & 70146 & 90842 \\ 75407 & 19716 & 10710 & 99562 & 60815 & 82473 & 51869 & 72201 \cdots. \end{matrix}$$

An investigation into the nature of the number ξ. The writer has been unable to discover any simple connection between ξ and other known constants. As to the nature of ξ, it is neither rational nor the root of a quadratic equation with rational coefficients, since its continued fraction is neither finite nor periodic. In what follows we show that ξ is not a root of a cubic equation with rational coefficients. We begin with

THEOREM 14. *Let a_ν and D_ν be the ν-th partial quotient and the denominator of the ν-th convergent of the continued fraction* (36). *Then $a_\nu > D_\nu$ for $\nu \leqq 4$.*

Remark. For $\nu = 1, 2, 3$, we have $a_\nu = D_\nu$.

Proof. The theorem is true for $\nu = 4$ since $a_4 = 34$, and $D_4 = 27$. If the theorem is true for $4 \leqq \nu < k$, we may prove it true for $\nu = k$ by showing that $a_{k-2}^{-1}(a_k - D_k)$ is positive. Using the fundamental recursion formula

$$D_{\mu+1} = a_\mu D_\mu + D_{\mu-1} \tag{37}$$

for $\mu = k-1, k-2, k-3$ we have

$$a_{k-2}^{-1}(a_k - D_k) = a_{k-2}^{-1}\left\{a_k - (a_{k-1}a_{k-2} + 1)D_{k-2} - a_{k-1}\frac{D_{k-2} - D_{k-4}}{a_{k-3}}\right\}. \tag{38}$$

If in (38) we introduce the hypothesis of the induction $D_{k-2} \leqq a_{k-2}$, we get

$$a_{k-2}^{-1}(a_k - D_k) \geqq a_k a_{k-2}^{-1} - (a_{k-1}a_{k-2} + 1) - a_{k-1}a_{k-3}^{-1}. \tag{39}$$

By (33) and (35) we may write (39) in the form

$$a_{k-2}^{-1}(a_k - D_k) \geqq n_{k-1} + n_{k-2} + 1 - (n_{k-1} - n_{k-2} + 1) - (n_{k-2} + n_{k-3} + 1)$$

$$= n_{k-2} - n_{k-3} - 1 = n_{k-3}^2 > 0.$$

Hence the induction is complete.

THEOREM 15. *The number ξ does not satisfy a cubic equation with rational coefficients.*

Proof. By Theorem 14 and the familiar inequality

$$\left|\xi - \frac{C_\nu}{D_\nu}\right| < \frac{1}{D_\nu D_{\nu+1}} = \frac{1}{D_\nu(D_\nu a_\nu + D_{\nu-1})} < \frac{1}{D_\nu^3},$$

it follows that the Diophantine inequality

$$\left|\xi - \frac{x}{y}\right| < \frac{1}{y^3} \tag{40}$$

has infinitely many solutions in integers (x, y). Now if ξ satisfied a cubic equation with rational coefficients, the cubic would be irreducible, since ξ is neither rational nor a root of a quadratic equation. By a theorem of Siegel[6] the inequality (40) would in this case have only a finite number of solutions in integers (x, y), contrary to fact.

To show that this type of argument cannot be used further to prove that ξ is not an algebraic number of degree > 3, we give

THEOREM 16. *If $\epsilon > 0$ and if c is a positive constant, no matter how large, the Diophantine inequality*

$$\left|\xi - \frac{x}{y}\right| < \frac{c}{y^{3+\epsilon}} \tag{41}$$

has only a finite number of solutions (x, y).

We first prove two other theorems.

THEOREM 17. *For every k the sequence*

$$\frac{D_k}{a_k}, \frac{D_{k+2}}{a_{k+2}}, \frac{D_{k+4}}{a_{k+4}}, \cdots$$

tends to a limit.

Proof. Since

$$\frac{D_{k+2N}}{a_{k+2N}} = \frac{D_k}{a_k} - \sum_{\lambda=0}^{N-1}\left\{\frac{D_{k+2\lambda}}{a_{k+2\lambda}} - \frac{D_{k+2(\lambda+1)}}{a_{k+2(\lambda+1)}}\right\},$$

[6] See Landau, *Vorlesungen über Zahlentheorie*, vol. 3, 1927, pp. 37–65.

it is sufficient to show that this series tends to a limit as $N \to \infty$. To examine its general term we replace $k + 2\lambda$ by ν for simplicity. Now

$$\frac{D_{\nu+2}}{a_{\nu+2}} = \frac{D_{\nu+1}a_{\nu+1} + D_\nu}{a_{\nu+2}} = \frac{D_\nu a_\nu a_{\nu+1} + D_{\nu-1}a_{\nu+1} + D_\nu}{a_{\nu+2}} = \frac{D_\nu(n_\nu^2 + 2)}{a_{\nu+2}} + \frac{D_{\nu-1}a_{\nu+1}}{a_{\nu+2}}$$

$$= \frac{D_\nu}{a_\nu}\frac{a_\nu}{a_{\nu+2}}(n_\nu^2 + 2) + \frac{D_{\nu-1}}{a_{\nu+2}^2}(n_{\nu+1}^2 + 1).$$

Hence the general term of the above series may be written

$$\frac{D_\nu}{a_\nu} - \frac{D_{\nu+2}}{a_{\nu+2}} = \frac{D_\nu}{a_\nu}\left(1 - \frac{a_\nu}{a_{\nu+2}}(n_\nu^2 + 2)\right) - \frac{D_{\nu-1}}{a_{\nu-1}}\frac{a_{\nu-1}}{a_{\nu+2}^2}(n_{\nu+1}^2 + 1).$$

By Theorem 14 it is sufficient to show that as ν runs over all numbers of the same parity as k the two infinite series

$$\sum\left(1 - \frac{a_\nu}{a_{\nu+2}}(n_\nu^2 + 2)\right), \qquad \sum \frac{a_{\nu-1}}{a_{\nu+2}^2}(n_{\nu+1}^2 + 1)$$

converge. The first of these may be written

$$\sum\left(1 - \frac{n_{\nu+1} - n_\nu + 1}{n_{\nu+1} + n_\nu + 1}\right) = \sum \frac{2n_\nu}{n_{\nu+1} + n_\nu + 1} < 2\sum\frac{1}{n_\nu},$$

a rapidly convergent series. As for the second series we have

$$\sum \frac{a_{\nu-1}}{a_\nu^2}\frac{n_{\nu+1}^2 + 1}{(n_{\nu+1} + n_\nu + 1)^2} < \sum\frac{a_{\nu-1}}{a_\nu^2} < \sum\frac{1}{a_\nu},$$

which also converges with rapidity. This completes the proof.

The two sequences

$$\frac{D_1}{a_1}, \frac{D_3}{a_3}, \frac{D_5}{a_5}, \cdots \qquad \text{and} \qquad \frac{D_2}{a_2}, \frac{D_4}{a_4}, \frac{D_6}{a_6}, \cdots$$

tend to different limits. In fact we find

$$\frac{D_6}{a_6} = .7898114735158, \qquad \frac{D_5}{a_5} = .9370558376,$$

$$\frac{D_8}{a_8} = .7898114728192, \qquad \frac{D_7}{a_7} = .9370280114.$$

These two limits we denote by R_0 and R_1. That is,

$$R_0 = \lim_{\nu\to\infty} D_{2\nu}/a_{2\nu} = .78981147 \cdots,$$

$$R_1 = \lim_{\nu\to\infty} D_{2\nu+1}/a_{2\nu+1} = .93702801 \cdots.$$

It can be proved without difficulty that the two sequences above are both strictly decreasing except for the fact that $\frac{D_1}{a_1} = \frac{D_3}{a_3}$. We are now in a position to prove

THEOREM 18. *If* $\nu \to \infty$, *then*

$$D_{2\nu}^3 \left| \xi - \frac{C_{2\nu}}{D_{2\nu}} \right| \to R_0 \qquad \text{and} \qquad D_{2\nu+1}^3 \left| \xi - \frac{C_{2\nu+1}}{D_{2\nu+1}} \right| \to R_1.$$

Proof. Since

$$\xi = \lim_{\nu=\infty} C_\nu / D_\nu,$$

we may write

$$\xi = \frac{C_\nu}{D_\nu} + \frac{C_{\nu+1}}{D_{\nu+1}} - \frac{C_\nu}{D_\nu} + \frac{C_{\nu+2}}{D_{\nu+2}} - \frac{C_{\nu+1}}{D_{\nu+1}} + \cdots.$$

Using the fundamental relation

$$C_\mu D_{\mu-1} - C_{\mu-1} D_\mu = (-1)^{\mu-1},$$

we obtain

$$D_\nu^3 \left| \xi - \frac{C_\nu}{D_\nu} \right| = \frac{D_\nu^2}{D_{\nu+1}} - \frac{D_\nu^3}{D_{\nu+1} D_{\nu+2}} + \frac{D_\nu^3}{D_{\nu+2} D_{\nu+3}} - \cdots. \tag{42}$$

The first term on the right may be shown to tend to R_0 or R_1 as follows:

$$\frac{D_{\nu+1}}{D_\nu^2} = \frac{D_\nu a_\nu + D_{\nu-1}}{D_\nu^2} = \frac{a_\nu}{D_\nu} + \frac{D_{\nu-1}}{D_\nu^2}.$$

As ν tends to infinity through integers of the same parity, $D_{\nu-1}/D_\nu^2$ tends rapidly to zero, while a_ν/D_ν tends to R_0^{-1} or R_1^{-1} according as the value of ν is even or odd by Theorem 17.

Each of the other terms of (42) tends to zero as $\nu \to \infty$ since for $\lambda > 0$

$$\frac{D_\nu^3}{D_{\nu+\lambda} D_{\nu+\lambda+1}} \leqq \frac{D_\nu^3}{D_{\nu+1} D_{\nu+2}} < \frac{D_\nu^3}{D_{\nu+1}^2} < \frac{D_\nu^3}{(D_\nu a_\nu)^2} < \frac{D_\nu}{a_\nu^2} < \frac{1}{a_\nu}$$

by Theorem 14. Hence the theorem is proved.

Theorem 16 now follows from Theorem 18 and from the fact that the convergents C_ν/D_ν are the fractions of best approximation to ξ.

We conclude with a theorem concerning the above-mentioned limits R_0 and R_1.

THEOREM 19.

$$R_0 R_1 = \frac{1}{1 + \xi^2}.$$

For the proof of this theorem we need the following result of interest in itself.

THEOREM 20. *If* C_ν/D_ν *is the* ν*-th convergent of* (36), *then*

$$C_\nu C_{\nu+1} + D_\nu D_{\nu+1} = n_{\nu+1}.$$

Proof. Let $A_{\nu+1}$ and $B_{\nu+1}$ be the numerator and denominator of the $(\nu + 1)$-st convergent $\sigma_{\nu+1}$ of the continued cotangent (8) defining ξ, or, what is the same, the numerator and denominator of the ν-th convergent of the continued fraction (31). From the theory of irregular continued fractions we have in view of (32) the following recursion formulas for the A's and the B's:

$$A_{\nu+1} = (n_{\nu-1}^2 + 1)(A_\nu + A_{\nu-1}), \tag{43}$$

$$B_{\nu+1} = (n_{\nu-1}^2 + 1)(B_\nu + B_{\nu-1}). \tag{44}$$

We now prove that

$$A_{\nu+1} = C_\nu a_\nu a_{\nu-1} a_{\nu-2} \cdots a_0 , \tag{45}$$

$$B_{\nu+1} = D_\nu a_\nu a_{\nu-1} a_{\nu-2} \cdots a_0 . \tag{46}$$

In fact, (45) is true for $\nu = 0$, since $A_1 = n_0 = 0$ and $C_0 = 0$. If (45) is true for $\nu < \mu$ we may prove it for $\nu = \mu$ as follows. By (43) and the hypothesis of induction

$$\begin{aligned} A_{\mu+1} &= (n_{\mu-1}^2 + 1)(C_{\mu-1} a_{\mu-1} a_{\mu-2} \cdots a_0 + C_{\mu-2} a_{\mu-2} \cdots a_0) \\ &= (n_{\mu-1}^2 + 1) a_{\mu-2} a_{\mu-3} \cdots a_0 (C_{\mu-1} a_{\mu-1} + C_{\mu-2}) = a_\mu a_{\mu-1} a_{\mu-2} \cdots a_0 \, C_\mu \end{aligned}$$

by (35). (46) is established in the same way. By Theorem 11

$$A_{\nu+1} A_{\nu+2} + B_{\nu+1} B_{\nu+2} = n_{\nu+1}(n_\nu^2 + 1)(n_{\nu-1}^2 + 1) \cdots (n_0^2 + 1), \tag{47}$$

while by (45) and (46)

$$\begin{aligned} A_{\nu+1} A_{\nu+2} + B_{\nu+1} B_{\nu+2} &= (C_\nu C_{\nu+1} + D_\nu D_{\nu+1})(a_{\nu+1} a_\nu)(a_\nu a_{\nu-1}) \cdots (a_1 a_0) a_0 \\ &= (C_\nu C_{\nu+1} + D_\nu D_{\nu+1})(n_\nu^2 + 1)(n_{\nu-1}^2 + 1) \cdots (n_0^2 + 1) a_0 . \end{aligned}$$

If we compare this with (47), the theorem is seen to follow from $a_0 \doteq 1$. In the same way Theorem 10 yields

THEOREM 21.

$$C_\nu^2 + D_\nu^2 = a_{\nu+1} .$$

Theorem 19 is now a simple consequence of Theorem 20. In fact, we have by definition of R_0 and R_1

$$\begin{aligned} R_0 R_1(\xi^2 + 1) &= \lim_{\nu \to \infty} \left(\frac{D_\nu}{a_\nu} \frac{D_{\nu+1}}{a_{\nu+1}} \left\{ \frac{C_\nu}{D_\nu} \frac{C_{\nu+1}}{D_{\nu+1}} + 1 \right\} \right) \\ &= \lim_{\nu \to \infty} \left(\frac{C_\nu C_{\nu+1} + D_\nu D_{\nu+1}}{n_{\nu+1} - n_\nu} \right) = \lim_{\nu \to \infty} \frac{n_{\nu+1}}{n_{\nu+1} - n_\nu} = 1. \end{aligned}$$

Regular continued cotangents of familiar constants. The converse problem of discovering a law enjoyed by the partial cotangents of the regular continued

cotangent expansion of a familiar constant appears to be even more difficult than in continued fractions. There are no periodic regular continued cotangents in view of (4). In fact, a periodic continued cotangent would not converge. Hence equation (22) cannot be used as in continued fractions to study the roots of a quadratic equation with rational coefficients. Furthermore, it is practically impossible to find more than 6 or 8 partial cotangents of a given irrational number. By Theorem 12, ten terms of the continued cotangent expansion of a number x between 10 and 11 would give x correctly to more than 1000 decimal places, 20 terms would give more than a million digits. This dependence of the continued cotangent expansion upon the "size" of x is brought out more sharply by the fact that two numbers x_1 and x_2 which merely differ by an integer may have widely different continued cotangent expansions while their continued fraction expansions are essentially the same. Thus, for example, $13/25 = \cot(\operatorname{arc\,cot} 0 - \operatorname{arc\,cot} 1 + \operatorname{arc\,cot} 3 - \operatorname{arc\,cot} 44)$, while $5 + (13/25) = \cot(\operatorname{arc\,cot} 5 - \operatorname{arc\,cot} 55)$.

The writer has been unable to discover any combination of familiar constants whose regular continued cotangent expansion is in any way predictable; that is, we have found nothing comparable with

$$\frac{3-e}{e-1} = \frac{1|}{|6} + \frac{1|}{|10} + \frac{1|}{|14} + \cdots + \frac{1|}{|4\nu+2} + \cdots$$

or with the irregular continued cotangent

$$2 + \sqrt{2} = \cot(\operatorname{arc\,cot} 3 + \operatorname{arc\,cot} 17 + \operatorname{arc\,cot} 99 + \operatorname{arc\,cot} 577 + \cdots)$$

whose partial cotangents satisfy the difference equation $n_{\nu+1} = 6n_\nu - n_{\nu-1}$. The continued cotangents for $\sqrt{2}$, π and e begin as follows:

$$\sqrt{2} = \cot(\operatorname{arc\,cot} 1 - \operatorname{arc\,cot} 5 + \operatorname{arc\,cot} 36 - \operatorname{arc\,cot} 3406 \\ + \operatorname{arc\,cot} 14694817 - \operatorname{arc\,cot} 727050997716715 + \cdots),$$

$$\pi = \cot(\operatorname{arc\,cot} 3 - \operatorname{arc\,cot} 73 + \operatorname{arc\,cot} 8599 - \operatorname{arc\,cot} 400091364 + \cdots),$$

$$e = \cot(\operatorname{arc\,cot} 2 - \operatorname{arc\,cot} 8 + \operatorname{arc\,cot} 75 - \operatorname{arc\,cot} 8949 \\ + \operatorname{arc\,cot} 11964723 \cdots).$$

Although this paper is concerned with developing the general properties of regular continued cotangents, the reader cannot have failed to notice that many of the theorems have number-theoretic implications. The applications of the above theory to Diophantine analysis will be given in another paper.

An interesting generalization of the regular continued cotangent is an expansion of the form

$$\cot \sum_{\nu=0} c_\nu \operatorname{arc\,cot} n_\nu, \tag{48}$$

in which c_ν are ± 1 and the n_ν satisfy certain inequalities. This is called a "semi-regular" continued cotangent and has many properties in common with the semi-regular continued fraction

$$q_0 \pm \frac{1|}{|q_1} \pm \frac{1|}{|q_2} \pm \cdots .$$

A discussion of semi-regular continued cotangents will appear later. However, if the coefficients c_ν of (48) are unrestricted, the analogy with continued fractions breaks down.

LEHIGH UNIVERSITY.

CHANGE OF DIMENSION IN SEQUENCE TRANSFORMATIONS

By Hugh J. Hamilton

In[1] H_1 were derived conditions on the $2n$-dimensional matrix of complex numbers $\| a_{m_1 m_2 \cdots m_n k_1 k_2 \cdots k_n} \|$ (n arbitrary) necessary and sufficient for the transforming of all n-dimensional sequences $\{s_{k_1 k_2 \cdots k_n}\}$ of class U into n-dimensional sequences $\{\sigma_{m_1 m_2 \cdots m_n}\}$ of class V, by means of the relation

$$\sigma_{m_1 m_2 \cdots m_n} = \sum_{k_1, k_2, \cdots, k_n = 1}^{\infty} a_{m_1 m_2 \cdots m_n k_1 k_2 \cdots k_n} s_{k_1 k_2 \cdots k_n},$$

for each U and each V of a set of 16 classes of sequences of complex numbers. It is the purpose of this note to point out that no use is made, in the course of the proofs in H_1, of equality between the dimension of the s-sequence and that of the σ-sequence, and consequently that the results, if properly interpreted, are quite valid for $(n + l)$-dimensional matrices $\| a_{m_1 m_2 \cdots m_l k_1 k_2 \cdots k_n} \|$ associated with transformations of the type

$$\sigma_{m_1 m_2 \cdots m_l} = \sum_{k_1, k_2, \cdots, k_n = 1}^{\infty} a_{m_1 m_2 \cdots m_l k_1 k_2 \cdots k_n} s_{k_1 k_2 \cdots k_n},$$

where l and n are independently arbitrary.

The only alterations which need to be made in H_1 to permit this interpretation follow.

Page 29.

Line 1. Replace "n-tuple" by "multiple".

Line 4. Replace n by l.

Line 5. Replace "another such, homologous to m", by "an ordered set of n positive, integral variables".

Line 16. Replace m^1, m^2 by k^1, k^2.

Line 18. Replace "The corresponding" by "A single".

Page 30.

Line 15. Before $\pi(m)$, insert "$l = n$, and"; and before "then", insert "μ being the homologue in m of κ in k".

Line 23. Before "under", insert "and $l = n$".

Line 26. Before "under", insert "and $l = n$".

Page 31.

Line 19. Before "and", insert "in all, $l = n$".

Received November 29, 1937; presented to the American Mathematical Society, April 16, 1938.

[1] In this note H_1 will denote the paper, *Transformations of multiple sequences*, by Hamilton, this Journal, vol. 2(1936), pp. 29–60.

Page 33.

Line 14. Replace $n(2^n - 2)$ by $n(2^l - 2)$.

As an application of this interpretation, consider the problem of stating conditions on the 3-dimensional matrix $\| a_{\nu ij} \|$, necessary and sufficient that the simple sequence $\{\sigma_\nu\}$ with $\sigma_\nu \equiv \sum_{i,j=1}^{\infty} a_{\nu ij} s_{ij}$ be defined for each ν and converge to $\lim_{i,j\to\infty} s_{ij}$ whenever the double sequence $\{s_{ij}\}$ is (A) boundedly convergent, (B) convergent. These can be read off (pp. 59–60 of H_1) as:

(A) (1) $$\sum_{i,j=1}^{\infty} | a_{\nu ij} | < A \qquad (\nu = 1, 2, \cdots);$$

(2) $$\lim_{\nu} \sum_{i,j=1}^{\infty} a_{\nu ij} = 1;$$

(3) $$\lim_{\nu} \sum_{i=1}^{\infty} | a_{\nu ij} | = 0 \qquad (j = 1, 2, \cdots),$$
$$\lim_{\nu} \sum_{j=1}^{\infty} | a_{\nu ij} | = 0 \qquad (i = 1, 2, \cdots).$$

(B) (1) and (2) same as (1) and (2) under (A), respectively;

(3) $$a_{\nu ij} = 0 \quad \text{for } i > C_j \qquad (\nu = 1, 2, \cdots; j = 1, 2, \cdots),$$
$$a_{\nu ij} = 0 \quad \text{for } j > C_i \qquad (\nu = 1, 2, \cdots; i = 1, 2, \cdots);$$

(4) $$\lim_{\nu} a_{\nu ij} = 0 \qquad (i, j = 1, 2, \cdots).$$

Fraleigh[2] has derived conditions on the matrix $\| a_{\nu ij} \|$ $[a_{\nu ij} = 0$ for $\max(i,j) > \nu]$ necessary and sufficient that $\lim_{\nu} \sigma_\nu = uv$, where $s_{ij} = u_i v_j$ $(i, j = 1, 2, \cdots)$, whenever $\lim_{i} u_i \equiv u$ and $\lim_{j} v_j \equiv v$ exist. Here $\{s_{ij}\}$ is a regularly convergent sequence, though of a special type, which is sometimes termed "factorable".

Now referring in H_1 as here revised to the conditions necessary and sufficient for the general transformation RC $\rightarrow$ BC regularly (pp. 59–60), with $l = 1$ and $n = 2$, as well as to the derivations of these conditions (which for d_3, $\bar{d}_1$, and $\bar{d}_2$ involve only factorable sequences), it will be found that the necessity of all of Fraleigh's conditions, excepting those difficult ones involving boundedness, and the sufficiency of the set in Fraleigh's Theorem III[3] are implicitly derived in H_1 under the present interpretation.

POMONA COLLEGE.

[2] P. A. Fraleigh, *Regular bilinear transformations of sequences*, American Journal of Mathematics, vol. 53 (1931), pp. 697–709.

[3] Fraleigh, loc. cit., p. 703.

A GENERALIZATION OF MULTIPLE SEQUENCE TRANSFORMATIONS

By Hugh J. Hamilton

1. **Introduction and statement of purpose.** It is the purpose of the present paper[1] to extend the definitions of the classes of complex multiple sequences $\{s_{k^{(1)},k^{(2)},\dots,k^{(n)}}\}$ considered in H_1 so that they may have meaning for complex functionals f of l variables $x^{(1)}, x^{(2)}, \cdots, x^{(l)}$, where, for $\nu = 1, 2, \cdots, l$, $x^{(\nu)}$ is the general element of an aggregate $\mathfrak{E}^{(\nu)}$ of considerable generality, interpretable, in particular, either as the positive integers or as the points of a Euclidean space; and to derive, by means of the results of $H_{1,2}$, conditions on the n-dimensional matrix $\|f_{k^{(1)},k^{(2)},\dots,k^{(n)}}\|$ of complex functionals of the l variables $x^{(1)}, x^{(2)}, \cdots, x^{(l)}$, necessary and sufficient for the various sequence class-to-functional class transformations under the relation

$$F(x^{(1)}, x^{(2)}, \cdots, x^{(l)}) \equiv \sum_{k^{(1)},k^{(2)},\dots,k^{(n)}=1}^{\infty} f_{k^{(1)},k^{(2)},\dots,k^{(n)}}(x^{(1)}, x^{(2)}, \cdots, x^{(l)}) s_{k^{(1)},k^{(2)},\dots,k^{(n)}}$$

analogous to the sequence class-to-sequence class transformations considered in $H_{1,2}$.

In order to clarify ideas, consider the solution of the problem: to find conditions on the linear matrix of complex functionals $\|f_i\|$ necessary and sufficient that $F(t) \equiv \sum_{i=1}^{\infty} f_i(t)s_i$ exist finite for each t in $(0, \infty)$ and converge to 0 as t tends to infinity, whenever $\{s_i\}$ is a null sequence of complex numbers.

Denote by S the class of non-negative, real sequences $\{t_j\}$ for which $\lim_j t_j = \infty$. Now $\lim_{t\to\infty} F(t) = 0$ if and only if $\lim_j F(t_j) = 0$ for each $\{t_j\} \in S$. Hence the desired conditions are simply that

$$(\alpha) \qquad \sum_{i=1}^{\infty} |f_i(t_j)| < B(\{t_j\}) \text{ for each } j \qquad (\text{each } \{t_j\} \in S);$$

$$(\beta) \qquad \lim_j f_i(t_j) = 0 \text{ for each } i \qquad (\text{each } \{t_j\} \in S),$$

as follows from well-known theorems of the Silverman-Toeplitz type (or from $H_{1,2}$); or, more elegantly, that

Received November 29, 1937; presented to the American Mathematical Society, April 16, 1938.

[1] This paper assumes familiarity with the contents of the preceding note, *Change of dimension in sequence transformations*; hence also with the paper therein cited as H_1. The paper H_1, as revised in the note, will be referred to as $H_{1,2}$.

$$(\alpha)\qquad \sum_{i=1}^{\infty} |f_i(t)| < \infty \text{ for all } t \text{ in } (0, \infty),$$
$$\sum_{i=1}^{\infty} |f_i(t)| < B \text{ for all } t > A;$$

$$(\beta)\qquad \lim_{t\to\infty} f_i(t) = 0 \text{ for each } i.$$

Here $l = n = 1$, $\mathfrak{C}^{(1)}$ is the set of non-negative real numbers, and the class of complex functionals f defined over $\mathfrak{C}^{(1)}$ for which $\lim_{t\to\infty} f(t) = 0$ would be called *null-convergent*. The statement and solution of this problem are a very particular specialization of the results of this paper. They have been displayed at this time and thus independently with the desire to make somewhat apparent at the outset the significance of the general problem, and to suggest reasons for the plan of approach used in its treatment.

In §2 are given definitions of most of the concepts and symbols which are to be used; §3 provides alternative definitions of certain classes of functionals introduced in §2; a list of conditions on the functionals $f_{k^{(1)}, k^{(2)}, \ldots, k^{(n)}}$ is made in §4; §§5 and 6 contain tabulations of sets of conditions necessary and sufficient for the various transformations; in §7 a few applications are indicated.

2. **Symbolism and definitions.** The letters $m, m^1, m^2, r^1, k, k^1, k^2, k^3, k^4, \kappa, \pi$, etc. and notations involving them, will have the meanings assigned to them in $\mathrm{H}_{1,2}$, n and l being independently arbitrary, but fixed, positive integers. Other symbols are to have the following meanings (each definition involving ν to apply, unless otherwise stated, for $\nu = 1, 2, \cdots, l$): $\mathfrak{C}^{(\nu)}$, an infinite aggregate of elements of any sort; $x^{(\nu)}$, the general element of $\mathfrak{C}^{(\nu)}$; x, the general ordered l-tuple $(x^{(1)}, x^{(2)}, \cdots, x^{(l)})$; $\mathfrak{C}$, the aggregate of x; x^1 and x^2, the generic ordered, proper, conjugate subsets of x; $y^{(\nu)}$, y, y^1, and y^2, a fixed element or set, of type $x^{(\nu)}$, x, x^1, and x^2, respectively; $x_i^{(\nu)}$, x_i, x_i^1, and x_i^2, the i-th in an infinite sequence of such elements or sets, respectively; $x{:}m$, the general set of form $(x^{(1)}_{m^{(1)}}, x^{(2)}_{m^{(2)}}, \cdots, x^{(l)}_{m^{(l)}})$, where $m^{(1)}, m^{(2)}, \cdots, m^{(l)}$ are the elements of m, in this order; $x^1{:}m^1$ and $x^2{:}m^2$ have analogous meanings.

$\mathfrak{S}^{(\nu)}$ is the general sequence $\{x_i^{(\nu)}\}$; $\mathfrak{S}$, the general l-tuple sequence of form $\{x{:}m\}$; $\mathfrak{S}^1$ and $\mathfrak{S}^2$, the general multiple sequences of form $\{x^1{:}m^1\}$ and $\{x^2{:}m^2\}$, respectively; $\mathfrak{R}_2$, $\mathfrak{R}_2^1$, and $\mathfrak{R}_2^2$, the classes of *all* sequences $\mathfrak{S}$, $\mathfrak{S}^1$, and $\mathfrak{S}^2$, respectively; $\{\mathfrak{C}_j^{(\nu)}\}$, an arbitrary, but fixed, infinite sequence of non-null sets for which $\mathfrak{C}^{(\nu)} \equiv \mathfrak{C}_1^{(\nu)} \supseteq \mathfrak{C}_2^{(\nu)} \supseteq \cdots$ and $\prod_{j=1}^{\infty} \mathfrak{C}_j^{(\nu)}$ is null; $h^{(\nu)}$, the functional over $\mathfrak{C}^{(\nu)}$ for which $h^{(\nu)}(x^{(\nu)})$ is the greatest integer j such that $x^{(\nu)} \in \mathfrak{C}_j^{(\nu)}$; $g^{(\nu)}$, an arbitrary, but fixed, real functional over $\mathfrak{C}^{(\nu)}$ for which $h^{(\nu)}(x^{(\nu)}) > M$ implies that $g^{(\nu)}(x^{(\nu)}) > N(M)$, where $\lim_{M\to\infty} N(M) = \infty$, and for which $g^{(\nu)}(x^{(\nu)}) > N$ implies that $h^{(\nu)}(x^{(\nu)}) > M(N)$, where $\lim_{N\to\infty} M(N) = \infty$; $g(x)$, the ordered set of numbers

$(g^{(1)}(x^{(1)}),\ g^{(2)}(x^{(2)}),\ \cdots,\ g^{(l)}(x^{(l)}))$; $g^1(x^1)$ and $g^2(x^2)$, the ordered sets defined analogously with respect to x^1 and x^2, respectively; $g^1(x^1) > D$, $g(x) \to \infty$, etc., the set of properties corresponding, for each separate $g^{(\nu)}(x^{(\nu)})$ involved; $\mathfrak{K}_1$, $\mathfrak{K}_1^1$, and $\mathfrak{K}_1^2$, the classes of all sequences $\mathfrak{S}$, $\mathfrak{S}^1$, and $\mathfrak{S}^2$, respectively, for which $\lim\limits_{m\to\infty} g(x:m) = \infty$, $\lim\limits_{m^1\to\infty} g^1(x^1:m^1) = \infty$, and $\lim\limits_{m^2\to\infty} g^2(x^2:m^2) = \infty$, respectively.

The $\mathfrak{S}$ *associated with* a particular $\{x_i\}$ is that $\mathfrak{S}$ for which $\mathfrak{S}^{(\nu)}$ is the sequence $\{x_i^{(\nu)}\}$; *the* $\mathfrak{S}^1$ [*or* $\mathfrak{S}^2$] *associated with* a particular $\{x_i^1\}$ [or $\{x_i^2\}$], that $\mathfrak{S}^1$ [or $\mathfrak{S}^2$] for which $\mathfrak{S}^{(\nu)}$ is the sequence $\{x_i^{(\nu)}\}$ for each ν involved; $\{s_k\}$, the general n-tuple sequence of complex numbers; s, $\lim\limits_{k\to\infty} s_k$ when this limit exists; f, the general complex functional over $\mathfrak{C}$; $\|f_k\|$, the n-tuple matrix of general complex functionals over $\mathfrak{C}$; F, the complex functional over $\mathfrak{C}$ defined by $F(x) \equiv \sum\limits_{k=1}^{\infty} f_k(x)s_k$, whenever this symbol has meaning.

It is hoped that such a consistency has been observed in defining these symbols as to facilitate keeping in mind their meanings. Certainly, use of the customary multiple sub- and superscripts would have both lengthened the paper greatly and hidden the essential simplicity of the proofs.

When, in any l- or n-tuple represented by a single letter, certain of the elements are to be considered as having special properties, this fact will be indicated in parentheses (or brackets, etc.), and subsequently parentheses (or brackets, etc.) will enclose the letter involved. Thus, "with $(x^1 = y^1)$ and $(k^3 = p^3)$, $\lim\limits_{g^2(x^2)\to\infty} \sum\limits_{k^4=1}^{\infty} F_{(k)}((x)) = L_{p^3}(y^1)$". A convention so established will last throughout, and only through, the paragraph in which it is set up.

Except when, by the nature of the situation, doing so would be obviously absurd, all relations involving subsets of k or of x are to be understood to imply the set of such relations for all possible choices of the subsets with respect to position and dimension.

Now the terms abbreviated by the symbols *e*, *ub*, *c*, *urc*, *rc*, *cn*, *urcn*, *rcn*, *urcrn*, *rcurn*, *rcrn*; *b*, *bc*, *burc*, *bcn*, *burcn*, and *burcrn* are, with their meanings for n-tuple sequences, indicated on pp. 30–31 of H_1. The symbols *brc*, *brcn*, *brcurn*, and *brcrn* abbreviate, respectively, *bounded regularly convergent*, *bounded regularly convergent null*, *bounded regularly convergent ultimately row null*, and *bounded regularly convergent row null*, and these terms indicate their own meanings. (As pointed out on p. 33 of H_1, each rc multiple sequence is b; hence the terms brc, brcn, brcurn, and brcrn are equivalent to, respectively, simply rc, rcn, rcurn, and rcrn, when applied to sequences. However, as will appear, this convenient property does not carry over to f in general, and it is for this reason that the four additional terms are introduced.)

The extended definitions of the twenty-one terms will now be given. With respect to the sequences of sets $\{\mathfrak{C}_j^{(\nu)}\}$, any complex functional f over $\mathfrak{C}$ will be said to be: *e*, if defined for all x; *ub*, if there exists a number Q such that $f(x)$

is bounded for all[2] $g(x) > Q$; *c*, if $\lim_{g(x)\to\infty} f(x) \equiv \sigma$ (read "principal limit") exists and is finite; *urc*, if c and if there exists a number $\bar{Q}$ such that $\lim_{g^1(x^1)\to\infty} f(x) \equiv \sigma(x^2)$ (read "row limit") exists and is finite for all $g^2(x^2) > \bar{Q}$; *rc*, if $\lim_{g^1(x^1)\to\infty} f(x) \equiv \sigma(x^2)$ exists and is finite for all x^2; *cn*, *urcn*, and *rcn*, if, respectively, c, urc, and rc, with $\sigma = 0$; *urcrn* and *rcurn* if, respectively, urc and rc with $\sigma(x^2) = 0$ for all $g^2(x^2)$ greater than some number $\bar{\bar{Q}}$; *rcrn*, if rc with $\sigma(x^2) = 0$ for all x^2; *b*, if $f(x)$ is bounded over $\mathfrak{E}$; *bc*, *burc*, *brc*, *bcn*, *burcn*, *brcn*, *burcrn*, *brcurn*, and *brcrn*, if both b and, respectively, c, urc, and so on. The notations E, UB, etc., will denote the classes of functionals f which are e, ub, etc., respectively.

That these definitions constitute a true generalization of those made for sequences in H_1 is seen by taking, for $\nu = 1, 2, \cdots, l$, $\mathfrak{E}^{(\nu)}$ to be the aggregate of the positive integers, $\mathfrak{E}_j^{(\nu)}$ the set of integers not less than j ($j = 1, 2, \cdots$), and $g^{(\nu)}(x^{(\nu)}) \equiv x^{(\nu)}$ for each $x^{(\nu)}$. Throughout this paper, classification of multiple sequences will be assumed to be made with respect to $\mathfrak{E}^{(\nu)}$ and $\mathfrak{E}_j^{(\nu)}$ as thus defined.

3. **Equivalent definitions of the several classes of functionals.** That f be of any class U of the set above which does not explicitly involve boundedness is equivalent to the condition that the sequence $\{f(x:m)\}$ be of class U for each $\mathfrak{S} \in \mathfrak{K}_1$. That the first condition implies the second is in each case obvious. That the second implies the first will now be proved for three typical cases.

THEOREM. *If* $\{f(x:m)\} \in E$ *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $f \in E$.

Proof. Each x is an element of some $\mathfrak{S} \in \mathfrak{K}_1$.

THEOREM. *If* $\{f(x:m)\} \in UB$ *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $f \in UB$.

Proof. By denial of the conclusion, there exists a sequence $\{x_i\}$ for which $\lim_i g(x_i) = \infty$ and $\{f(x_i)\}$ is unbounded. The associated $\mathfrak{S}$ belongs to $\mathfrak{K}_1$, but $\{f(x:m)\} \epsilon' UB$.[3]

THEOREM. *If* $\{f(x:m)\} \in URC$ *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $f \in URC$.

Proof. By denial of the conclusion, there exist a sequence $\{x_i^1\}$ and a set of sequences $\{{}_ix_j^2\}$ for which $\lim_i g^1(x_i^1) = \infty$ and $\lim_j g^2({}_ix_j^2) = \infty$ ($i = 1, 2, \cdots$), and, with ($x^1 = x_i^1$ and $x^2 = {}_ix_j^2$), for $i = 1, 2, \cdots$, $\lim_j f((x))$ does not exist. Let the sequence $\{x_j^2\}$ be so constructed of elements of the sequences $\{{}_ix_j^2\}$ that $\lim_j g^2(x_j^2) = \infty$ and, with $[x^1 = x_i^1$ and $x^2 = x_j^2]$, for $i = 1, 2, \cdots$, $\lim_j f([x])$ fails to exist. Now the sequences $\mathfrak{S}^1$ and $\mathfrak{S}^2$ associated with $\{x_i^1\}$ and $\{x_j^2\}$, respectively, belong to $\mathfrak{K}_1^1$ and $\mathfrak{K}_1^2$, respectively, so that their union, $\mathfrak{S}$, belongs to $\mathfrak{K}_1$. But $\{f(x:m)\} \epsilon' URC$.

That $f \in B$ is equivalent to the condition that $\{f(x:m)\} \in B$ for each $\mathfrak{S} \in \mathfrak{K}_2$. This is clear.

[2] That the definitions about to be given are equivalent for each fixed functional $g^{(\nu)}$ ($\nu = 1, 2, \cdots, l$) satisfying the conditions imposed above will be obvious.

[3] The notation ϵ' means "is not an element of."

Now, given any $\mathfrak{S}$, for the matrix of numbers $\| f_k(x:m) \|$ NS[4] the sequence $\{F(x:m)\}$ be of a particular one of the classes, say V, whenever $\{s_k\}$ is of a particular one, say U, are provided by $H_{1,2}$. [Take $a_{mk} \equiv f_k(x:m)$.] Hence, in view of the above established equivalences, for $\| f_k \|$ NS $F \epsilon V$ whenever $\{s_k\} \epsilon U$ is that the set of conditions for the corresponding sequence class-to-sequence class transformation given in $H_{1,2}$ apply to $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_1$ in case boundedness is not explicitly demanded in V, and apply to $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_2$ in case $V = B$. Examination of the theorems in $H_{1,2}$ reveals that each of the conditions listed on pp. 35–37 of $H_{1,2}$ is one of some set of conditions[5] NS $U \rightarrow V$, where V does not explicitly involve boundedness, and that the conditions (c_1) and (c_2) alone appear in sets of NS $U \rightarrow B$, for each U and V of the 16 classes of sequences. The next section, then, will be devoted to simplification of the conditions that each of the conditions on pp. 35–37 of $H_{1,2}$, separately, hold for $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_1$, and that (c_1) and (c_2), separately, hold for $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_2$.

4. **Reduction of conditions.** A condition on $\| f_k \|$ which is equivalent to the application of any particular condition from the set on pp. 35–37 of $H_{1,2}$ to $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_1$ will be labeled by the capital of the letter labeling the condition in $H_{1,2}$. The conditions which are equivalent to the application of conditions (c_1) and (c_2) of $H_{1,2}$ to $\| f_k(x:m) \|$ for each $\mathfrak{S} \epsilon \mathfrak{K}_2$ will be labeled $(\bar{C}_1)$ and $(\bar{C}_2)$, respectively. The conditions are listed first, and justifications will follow.

E conditions.

(A_1)
$$\sum_{k=1}^{\infty} | f_k(x) | < \infty \qquad \text{(all } x).$$

(A_2) Let κ, λ be any two single elements of k, and k^2 those remaining. Then $f_k(x) = 0$ for $\lambda > C_\kappa(x)$ ($k^2 = 1, 2, \cdots$; all x; $\kappa = 1, 2, \cdots$).

UB conditions.

(B_1)
$$\sum_{k=1}^{\infty} | f_k(x) | < A \text{ for } g(x) > B.$$

(B_2) Let κ, λ be any two single elements of k, and k^2 those remaining. Then $f_k(x) = 0$ for $g(x), \lambda > C_\kappa$ ($k^2 = 1, 2, \cdots$; $\kappa = 1, 2, \cdots$).

Quasi-B conditions.[6]

$$\sum_{k=1}^{\infty} | f_k(x) | < \infty \qquad \text{(all } x),$$

[4] NS will abbreviate *condition(s) necessary and sufficient that.* Throughout the paper it will be demanded of every transformation that $F \epsilon E$.

[5] $U \rightarrow V$ [or $U, W, \cdots \rightarrow V$] means *each element of class* U [*or* $U + W + \cdots$] *be transformed into some element of class* V.

[6] *Quasi-B conditions,* in contradistinction to B conditions $(\bar{C}_1)$ and $(\bar{C}_2)$ below.

(C$_1$)
$$\sum_{k=1}^{\infty} |f_k(x)| < A \text{ for } g(x) > B,$$

$$\sum_{k=1}^{\infty} |f_k(x)| < A(x^2) \text{ for } g^1(x^1) > B(x^2) \qquad (\text{all } x^2).$$

(C$_2$) Let κ, λ be any two single elements of k, and k^2 those remaining. Then

$$\left.\begin{array}{ll} f_k(x) = 0 \text{ for } \lambda > C_\kappa(x) & (\text{all } x), \\ f_k(x) = 0 \text{ for } g(x), \lambda > C_\kappa, & \\ f_k(x) = 0 \text{ for } g^1(x^1), \lambda > C_\kappa(x^2) & (\text{all } x^2) \end{array}\right\} (k^2 = 1, 2, \cdots ; \kappa = 1, 2, \cdots).$$

C conditions.

(D$_1$)
$$\lim_{g(x)\to\infty} f_k(x) = a_k \qquad (k = 1, 2, \cdots).$$

(D$_2$)
$$\lim_{g(x)\to\infty} \sum_{k^2=1}^{\infty} f_k(x) = L_{k^1} \qquad (k^1 = 1, 2, \cdots).$$

(D$_3$)
$$\lim_{g(x)\to\infty} \sum_{k=1}^{\infty} f_k(x) = L.$$

(D$_4$) There exist numbers a_k such that, if κ is any single element of k, and k^2 those remaining, then

$$\lim_{g(x)\to\infty} \sum_{k^2=1}^{\infty} |f_k(x) - a_k| = 0 \qquad (\kappa = 1, 2, \cdots).$$

(D$_5$) There exist numbers a_k such that

$$\lim_{g(x)\to\infty} \sum_{k=1}^{\infty} |f_k(x) - a_k| = 0.$$

CN conditions.

($\bar{\text{D}}_1$) (D$_1$), with $a_k = 0$ $(k = 1, 2, \cdots)$.

($\bar{\text{D}}_2$) (D$_2$), with $L_{k^1} = 0$ $(k^1 = 1, 2, \cdots)$.

($\bar{\text{D}}_3$) (D$_3$), with $L = 0$.

($\bar{\text{D}}_4$) (D$_4$), with $a_k = 0$ $(k = 1, 2, \cdots)$.

($\bar{\text{D}}_5$) (D$_5$), with $a_k = 0$ $(k = 1, 2, \cdots)$.

URC conditions.

(E$_1$)
$$\lim_{g^2(x^2)\to\infty} f_k(x) = a_k(x^1) \text{ for } g^1(x^1) > D \qquad (k = 1, 2, \cdots).$$

(E$_2^*$) Let κ be any single element of k^3, and k^{32} (which may in this case be null) those remaining. Then

$$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} f_k(x) = L_{k^3}(x^1) \text{ for } g^1(x^1) > E_\kappa \qquad (k^{32} = 1, 2, \cdots ; \kappa = 1, 2, \cdots).$$

(E₂) $$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} f_k(x) = L_{k^3}(x^1) \text{ for } g^1(x^1) > E \quad (k^3 = 1, 2, \cdots).$$

(E₃) $$\lim_{g^2(x^2)\to\infty} \sum_{k=1}^{\infty} f_k(x) = L(x^1) \text{ for } g^1(x^1) > F.$$

(E_4^*) There exist numbers $a_k(x^1)$ such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} |f_k(x) - a_k(x^1)| = 0 \text{ for } g^1(x^1) > G_\kappa \quad (\kappa = 1, 2, \cdots).$$

(E_4) There exist numbers $a_k(x^1)$ such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} |f_k(x) - a_k(x^1)| = 0 \text{ for } g^1(x^1) > G \quad (\kappa = 1, 2, \cdots).$$

(E_5) There exist numbers $a_k(x^1)$ such that

$$\lim_{g^2(x^2)\to\infty} \sum_{k=1}^{\infty} |f_k(x) - a_k(x^1)| = 0 \text{ for } g^1(x^1) > H.$$

URCRN conditions.

($\bar{E}_1$) (E_1), with $a_k(x^1) = 0$ for $g^1(x^1) > \bar{D}$ $(k = 1, 2, \cdots)$.

($\bar{E}_2^*$) (E_2^*), with $L_{k^3}(x^1) = 0$ for $g^1(x^1) > \bar{E}_\kappa$

$(k^{32} = 1, 2, \cdots; \kappa = 1, 2, \cdots)$.

($\bar{E}_2$) (E_2), with $L_{k^3}(x^1) = 0$ for $g^1(x^1) > \bar{E}$ $(k^3 = 1, 2, \cdots)$.

($\bar{E}_3$) (E_3), with $L(x^1) = 0$ for $g^1(x^1) > \bar{F}$.

($\bar{E}_4^*$) (E_4^*), with $a_k(x^1) = 0$ for $g^1(x^1) > \bar{G}_\kappa$

$(k^4 = 1, 2, \cdots; \kappa = 1, 2, \cdots)$.

($\bar{E}_4$) (E_4), with $a_k(x^1) = 0$ for $g^1(x^1) > \bar{G}$ $(k = 1, 2, \cdots)$.

($\bar{E}_5$) (E_5), with $a_k(x^1) = 0$ for $g^1(x^1) > \bar{H}$ $(k = 1, 2, \cdots)$.

RC conditions.

(F_1) $$\lim_{g^2(x^2)\to\infty} f_k(x) = a_k(x^1) \text{ for all } x^1 \quad (k = 1, 2, \cdots).$$

(F_2) $$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} f_k(x) = L_{k^3}(x^1) \text{ for all } x^1 \quad (k^3 = 1, 2, \cdots).$$

(F_3) $$\lim_{g^2(x^2)\to\infty} \sum_{k=1}^{\infty} f_k(x) = L(x^1) \text{ for all } x^1.$$

(F_4) There exist numbers $a_k(x^1)$ such that, if κ is any single element of k, and k^4 those remaining, then

$$\lim_{g^2(x^2)\to\infty} \sum_{k^4=1}^{\infty} |f_k(x) - a_k(x^1)| = 0 \text{ for all } x^1 \qquad (\kappa = 1, 2, \cdots).$$

(F_5) There exist numbers $a_k(x^1)$ such that

$$\lim_{g^2(x^2)\to\infty} \sum_{k=1}^{\infty} |f_k(x) - a_k(x^1)| = 0 \text{ for all } x^1.$$

RCRN conditions.

($\bar{F}_1$) (F_1), with $a_k(x^1) = 0$ for all x^1 $(k = 1, 2, \cdots)$.

($\bar{F}_2$) (F_2), with $L_{k^3}(x^1) = 0$ for all x^1 $(k^3 = 1, 2, \cdots)$.

($\bar{F}_3$) (F_3), with $L(x^1) = 0$ for all x^1.

($\bar{F}_4$) (F_4), with $a_k(x^1) = 0$ for all x^1 $(k = 1, 2, \cdots)$.

($\bar{F}_5$) (F_5), with $a_k(x^1) = 0$ for all x^1 $(k = 1, 2, \cdots)$.

B conditions.

($\bar{C}_1$)
$$\sum_{k=1}^{\infty} |f_k(x)| < A \qquad (\text{all } x).$$

($\bar{C}_2$) Let κ, λ be any two single elements of k, and k^2 those remaining. Then

$$f_k(x) = 0 \text{ for } \lambda > C_\kappa \qquad (k^2 = 1, 2, \cdots; \text{all } x; \kappa = 1, 2, \cdots).$$

That each of these conditions implies the corresponding condition in $H_{1,2}$ [with $a_{mk} \equiv f_k(x:m)$] for each $\mathfrak{S} \in \mathfrak{K}_\mu$, μ having its proper value, is obvious, excepting perhaps (C_1) and (C_2). Proofs of the fact for these two follow.

THEOREM. *If* $\| f_k \|$ *satisfies* (C_1), *then* $\| f_k(x:m) \|$ *satisfies* (c_1) *for each* $\mathfrak{S} \in \mathfrak{K}_1$.

Proof. Let $\mathfrak{S} \in \mathfrak{K}_1$. Define $B(R) \equiv 1 + \max [R, B(x^2:m^2)]$ over $g^2(x^2:m^2) \leqq R$ for all possible dimensions and positions of x^2 with respect to x; $B_i \equiv B(B_{i-1})$ for $i = 2, 3, \cdots, l$; $B_1 \equiv B$; $\bar{A} \equiv \max A(x^2:m^2)$ over $g^2(x^2:m^2) \leqq B_{l-1}$ for all possible dimensions and positions of x^2 with respect to x; $\bar{\bar{A}} \equiv 1 + \max \sum_{k=1}^{\infty} |f_k(x:m)|$ over $g(x:m) \leqq B_l$. Let now $x:m$ be an arbitrary element of $\mathfrak{S}$. If $g(x:m) > B$, then $\sum_{k=1}^{\infty} |f_k(x:m)| < A$. If $g(x:m) \leqq B_l$, then $\sum_{k=1}^{\infty} |f_k(x:m)| < \bar{\bar{A}}$. In the remaining case, there exist an i $(2 \leqq i \leqq l)$ and conjugate subsets $x^1:m^1$ and $x^2:m^2$ of $x:m$ for which $g^1(x^1:m^1) > B_i$ and $g^2(x^2:m^2) \leqq B_{i-1}$. Hence $\sum_{k=1}^{\infty} |f_k(x:m)| < \bar{A}$.

THEOREM. *If* $\| f_k \|$ *satisfies* (C_2), *then* $\| f_k(x:m) \|$ *satisfies* (c_2) *for each* $\mathfrak{S} \in \mathfrak{K}_1$.

Proof. Let $\mathfrak{S} \in \mathfrak{K}_1$, fix upon particular positions for κ, λ, and let the value of κ be fixed. Define $C_\kappa(R) \equiv 1 + \max [R, C_\kappa(x^2:m^2)]$ over $g^2(x^2:m^2) \leqq R$ for

all possible dimensions and positions of x^2 with respect to x; $C_{\kappa i} \equiv C_\kappa(C_{\kappa,i-1})$ for $i = 2, 3, \cdots, l$; $C_{\kappa 1} \equiv C_\kappa$; $\bar{C}_\kappa \equiv \max C_\kappa(x^2:m^2)$ over $g^2(x^2:m^2) \leqq C_{\kappa,l-1}$ for all possible dimensions and positions of x^2 with respect to x; $\bar{\bar{C}}_\kappa \equiv \max C_\kappa(x:m)$ over $g(x:m) \leqq C_{\kappa l}$. Let now $x:m$ be an arbitrary element of $\mathfrak{S}$. If $g(x:m) > C_\kappa$, then $f_k(x) = 0$ for $\lambda > C_\kappa$ ($k^2 = 1, 2, \cdots$). If $g(x:m) \leqq C_{\kappa l}$, then $f_k(x) = 0$ for $\lambda > \bar{\bar{C}}_\kappa$ ($k^2 = 1, 2, \cdots$). In the remaining case, there exist an i ($2 \leqq i \leqq l$) and conjugate subsets $x^1:m^1$ and $x^2:m^2$ of $x:m$ for which $g^1(x^1:m^1) > C_{\kappa i}$ and $g^2(x^2:m^2) \leqq C_{\kappa,i-1}$. Hence $f_k(x:m) = 0$ for $\lambda > \bar{C}_\kappa$ ($k^2 = 1, 2, \cdots$).

Conversely, given any particular condition from pp. 35–37 of $\mathrm{H}_{1,2}$, its application to $\| a_{mk} \|$, where $a_{mk} \equiv f_k(x:m)$, for each $\mathfrak{S} \in \mathfrak{K}_1$ implies the corresponding condition on $\| f_k \|$ from the present set [excluding $(\bar{\mathrm{C}}_1)$ and $(\bar{\mathrm{C}}_2)$]; and the application of (c_1) [or (c_2)] from $\mathrm{H}_{1,2}$ to $\| f_k(x:m) \|$ for each $\mathfrak{S} \in \mathfrak{K}_2$ implies $(\bar{\mathrm{C}}_1)$ [or $(\bar{\mathrm{C}}_2)$, respectively]. Certain typical proofs will now be given.

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (a_1) *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $\| f_k \|$ *satisfies* (A_1).

Proof. Each x is an element of some $\mathfrak{S} \in \mathfrak{K}_1$.

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (b_1) *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $\| f_k \|$ *satisfies* (B_1).

Proof. By denial of the conclusion, there exists a sequence $\{x_i\}$ for which $\lim_i g(x_i) = \infty$ and $\left\{ \sum_{k=1}^{\infty} | f_k(x_i) | \right\}$ is unbounded. The associated $\mathfrak{S}$ belongs to $\mathfrak{K}_1$, but $\| f_k(x:m) \|$ does not satisfy (b_1).

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (c_1) *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $\| f_k \|$ *satisfies* (C_1).

Proof. It suffices to treat the third part only. By denial of the conclusion for this part, there exist y^2 and a sequence $\{x_i^1\}$ for which $\lim_i g^1(x_i^1) = \infty$ and, with $(x^1 = x_i^1$ and $x^2 = y^2)$, the sequence $\left\{ \sum_{k=1}^{\infty} | f((x)) | \right\}$ is unbounded. Now the sequence $\mathfrak{S}^1$ associated with $\{x_i^1\}$ belongs to $\mathfrak{K}_1^1$, and y^2 is an element of some $\mathfrak{S}^2 \in \mathfrak{K}_1^2$. Hence $\mathfrak{S} \in \mathfrak{K}_1$, but $\| f_k(x:m) \|$ does not satisfy (c_1).

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (d_1) *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $\| f_k \|$ *satisfies* (D_1).

Proof. By denial of the conclusion, there exist k and a sequence $\{x_i\}$ for which $\lim_i g(x_i) = \infty$ and $\lim_i f_k(x_i)$ does not exist. The associated $\mathfrak{S}$ belongs to $\mathfrak{K}_1$, but $\| f_k(x:m) \|$ does not satisfy (d_1).

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (d_4) *for each* $\mathfrak{S} \in \mathfrak{K}_1$, *then* $\| f_k \|$ *satisfies* (D_4).

Proof. Let $\mathfrak{S} \in \mathfrak{K}_1$. Then there exist numbers $a_k(\mathfrak{S})$ such that

$$\lim_{m\to\infty} \sum_{k^2=1}^{\infty} | f_k(x:m) - a_k(\mathfrak{S}) | = 0 \qquad (\kappa = 1, 2, \cdots).$$

By denial of the conclusion of the theorem, then, for some π and $\{\bar{x}_i\}$ for which $\lim_i g(\bar{x}_i) = \infty$, with $(\kappa = \pi)$, it is not true that

$$\lim_i \sum_{k^2=1}^{\infty} | f_{(k)}(\bar{x}_i) - a_{(k)}(\mathfrak{S}) | = 0.$$

Let $\{x_i\}$ be the sequence with which $\mathfrak{S}$ is associated, and $\{\bar{x}_i\}$ the sequence obtained by alternating the elements of $\{x_i\}$ and $\{\bar{x}_i\}$. The associated $\bar{\bar{\mathfrak{S}}}$ belongs to $\mathfrak{K}_1$, but $\|f_k(\bar{\bar{x}}_i:m)\|$ does not satisfy (d_4).

THEOREM. *If* $\|f_k(x:m)\|$ *satisfies* (e_1) *for each* $\mathfrak{S} \epsilon \mathfrak{K}_1$, *then* $\|f_k\|$ *satisfies* (E_1).

Proof. By denial of the conclusion, there exist sequences $\{x_i^1\}$ and $\{k_i\}$ and a set of sequences $\{{}_ix_j^2\}$ for which $\lim_i g^1(x_i^1) = \infty$ and $\lim_j g^2({}_ix_j^2) = \infty$ ($i = 1, 2, \cdots$), such that, with ($x^1 = x_i^1$ and $x^2 = {}_ix_j^2$), for $i = 1, 2, \cdots$, $\lim_j f_{k_i}((x))$ does not exist. Let the sequence $\{x_j^2\}$ be so constructed of the elements of the sequences $\{{}_ix_j^2\}$ that $\lim_j g^2(x_j^2) = \infty$ and, with [$x^1 = x_i^1$ and $x^2 = x_j^2$], for $i = 1, 2, \cdots$, $\lim_j f_{k_i}([x])$ fails to exist. Now the sequences $\mathfrak{S}^1$ and $\mathfrak{S}^2$ associated with $\{x_i^1\}$ and $\{x_j^2\}$, respectively, belong to $\mathfrak{K}_1^1$ and $\mathfrak{K}_1^2$, respectively, so that $\mathfrak{S}$ belongs to $\mathfrak{K}_1$. But $\|f_k(x:m)\|$ does not satisfy (e_1).

THEOREM. *If* $\|f_k(x:m)\|$ *satisfies* (e_4^*) *for each* $\mathfrak{S} \epsilon \mathfrak{K}_1$, *then* $\|f_k\|$ *satisfies* (E_4^*).

Proof. Part I. Let $\mathfrak{S}^2 \epsilon \mathfrak{K}_1^2$. Then there exist numbers $a_k(x^1; \mathfrak{S}^2)$ such that, with ($x^2 = x^2:m^2$),

$$\lim_{m^2\to\infty} \sum_{k^4=1}^{\infty} |f_k((x)) - a_k(x^1; \mathfrak{S}^2)| = 0$$

for $g^1(x^1) > G'_\kappa$ ($\kappa = 1, 2, \cdots$). For otherwise there exist π and a sequence $\{x_i^1\}$ for which $\lim_i g^1(x_i^1) = \infty$ such that, with [$x^1 = x_i^1$ and $x^2 = x^2:m^2$] and ($\kappa = \pi$), there exist, for $i = 1, 2, \cdots$, no numbers $a_{(k)}(x_i^1; \mathfrak{S}^2)$ for which

$$\lim_{m^2\to\infty} \sum_{k^4=1}^{\infty} |f_{(k)}([x]) - a_{(k)}(x_i^1; \mathfrak{S}^2)| = 0,$$

so that, if $\mathfrak{S}^1$ is associated with $\{x_i^1\}$ and hence belongs to $\mathfrak{K}_1^1$, $\|f_k(x:m)\|$ does not satisfy (e_4^*).

Part II. By denial of the conclusion of the theorem, there exist ρ and a sequence $\{\bar{x}_i^1\}$ and a set of sequences $\{{}_i\bar{x}_j^2\}$ for which $\lim_i g^1(x_i^1) = \infty$ and $\lim_j g^2({}_i\bar{x}_j^2) = \infty$ ($i = 1, 2, \cdots$) such that, with ($x^1 = \bar{x}_i^1$ and $x^2 = {}_i\bar{x}_j^2$) and ($\kappa = \rho$), for $i = 1, 2, \cdots$, it is not true that

$$\lim_j \sum_{k^4=1}^{\infty} |f_{(k)}((x)) - a_{(k)}(\bar{x}_i^1; \mathfrak{S}^2)| = 0.$$

Let the sequence $\{\bar{x}_j^2\}$ be so constructed of the elements of the sequences $\{{}_i\bar{x}_j^2\}$ that $\lim_j g^2(\bar{x}_j^2) = \infty$ and, with [$x^1 = \bar{x}_i^1$ and $x^2 = \bar{x}_j^2$], for $i = 1, 2, \cdots$, it is not true that

$$\lim_j \sum_{k^4=1}^{\infty} |f_{(k)}([x]) - a_{(k)}(\bar{x}_i^1; \mathfrak{S}^2)| = 0.$$

Let $\{x_j^2\}$ be the sequence with which $\mathfrak{S}^2$ is associated, and $\{\bar{x}_j^2\}$ the sequence obtained by alternating the elements of $\{x_j^2\}$ and $\{\bar{x}_j^2\}$. The associated $\bar{\bar{\mathfrak{S}}}^2$ belongs to $\mathfrak{R}_1^2$, but does not satisfy the conclusion of Part I.

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (c_1) *for each* $\mathfrak{S} \in \mathfrak{R}_2$, *then* $\| f_k \|$ *satisfies* $(\bar{C}_1)$.

Proof. By denial of the conclusion, there exists a sequence $\{x_i\}$ for which $\left\{\sum_{k=1}^{\infty} | f_k(x_i) | \right\}$ is unbounded. The associated $\mathfrak{S}$ belongs to $\mathfrak{R}_2$, but $\| f_k(x:m) \|$ does not satisfy (c_1).

THEOREM. *If* $\| f_k(x:m) \|$ *satisfies* (c_2) *for each* $\mathfrak{S} \in \mathfrak{R}_2$, *then* $\| f_k \|$ *satisfies* $(\bar{C}_2)$.

Proof. By denial of the conclusion, for some π and sequences $\{k_i^2\}$, $\{x_i\}$, and $\{\lambda_i\}$ with $\lim_i \lambda_i = \infty$, with ($\kappa = \pi$, $\lambda = \lambda_i$, and $k^2 = k_i^2$), for $i = 1, 2, \cdots$, $f_{(k)}(x_i) \neq 0$. The associated $\mathfrak{S}$ belongs to $\mathfrak{R}_2$, but $\| f_k(x:m) \|$ does not satisfy (c_2).

5. **Tabulation of results.** In view of the remarks in the last paragraph of §3, NS $U \to V$ can be derived immediately from $H_{1,2}$ for each U of the 16 classes of sequences named in H_1, and each V of the 21 classes of functionals named here,[7] in the following way. If V does not explicitly involve boundedness, the conditions are simply those labeled by the capitals of the letters labeling the conditions for the corresponding sequence class-to-sequence class transformation treated in $H_{1,2}$. If $V = B$, the conditions are either $(\bar{C}_1)$ or $(\bar{C}_1)$ and $(\bar{C}_2)$, according as (c_1) or (c_1) and (c_2), respectively, are necessary and sufficient for the corresponding transformation to bounded sequences. If $V \neq B$ but explicitly involves boundedness, the conditions are those for the transformation for the corresponding V which does not involve boundedness, plus the conditions for $V = B$.

Since the results for the various transformations were not explicitly stated in $H_{1,2}$, and since the present results are of large generality, the condensed tabulation which follows seems warranted.

E transformations.

1. NS (B, BC, BCN, BURC, BURCN, BURCRN, RC, RCN, RCURN, RCRN) → E is (A_1).
2. NS (UB, C, CN, URC, URCN, URCRN) → E are (A_1) and (A_2).

UB transformations.

3. NS (B, BC, BCN, BURC, BURCN, BURCRN, RC, RCN, RCURN, RCRN) → UB are (A_1) and (B_1).
4. NS (UB, C, CN, URC, URCN, URCRN) → UB are (A_1), (A_2), (B_1), and (B_2).

C transformations.

5. NS RCRN → C are (A_1), (B_1), and (D_1).
6. NS (RCN, RCURN) → C are (A_1), (B_1), (D_1), and (D_2).
7. NS RC → C are (A_1), (B_1), (D_1), (D_2), and (D_3).

[7] That is, with existence of $F(x)$ for each x demanded.

8. NS (BCN, BURCN, BURCRN) → C are (A_1), (B_1), and (D_4).
9. NS (BC, BURC) → C are (A_1), (B_1), (D_3), and (D_4).
10. NS B → C are (A_1), (B_1), and (D_5).
11. NS (CN, URCN, URCRN) → C are (A_1), (A_2), (B_1), (B_2), and (D_1).
12. NS (C, URC) → C are (A_1), (A_2), (B_1), (B_2), (D_1), and (D_3).
13. NS UB → C are (A_1), (A_2), (B_1), (B_2), and (D_5).

For CN transformations, replace in the above set each C condition by the corresponding CN condition. (Here, in altering 10 and 13, (B_1) can be omitted.)

URC transformations.

14. NS RCRN → URC are (A_1), (B_1), (D_1), and (E_1).
15. NS RCURN → URC are (A_1), (B_1), (D_1), (D_2), (E_1), and (E_2^*).
16. NS RCN → URC are (A_1), (B_1), (D_1), (D_2), (E_1), and (E_2).
17. NS RC → URC are (A_1), (B_1), (D_1), (D_2), (D_3), (E_1), (E_2), and (E_3).
18. NS BURCRN → URC are (A_1), (B_1), (D_4), (E_1), and (E_4^*).
19. NS BURCN → URC are (A_1), (B_1), (D_4), (E_1), (E_2), and (E_4^*).
20. NS BURC → URC are (A_1), (B_1), (D_3), (D_4), (E_1), (E_2), (E_3), and (E_4^*).
21. NS BCN → URC are (A_1), (B_1), (D_4), and (E_4).
22. NS BC → URC are (A_1), (B_1), (D_3), (D_4), (E_3), and (E_4).
23. NS B → URC are (A_1), (B_1), (D_5), and (E_5).
24. NS URCRN → URC are (A_1), (A_2), (B_1), (B_2), (D_1), and (E_1).
25. NS URCN → URC are (A_1), (A_2), (B_1), (B_2), (D_1), (E_1), and (E_2).
26. NS URC → URC are (A_1), (A_2), (B_1), (B_2), (D_1), (D_3), (E_1), (E_2), and (E_3).
27. NS CN → URC are (A_1), (A_2), (B_1), (B_2), (D_1), and (E_4).
28. NS C → URC are (A_1), (A_2), (B_1), (B_2), (D_1), (D_3), (E_3), and (E_4).
29. NS UB → URC are (A_1), (A_2), (B_1), (B_2), (D_5), and (E_5).

For URCN transformations, replace in the above set each C condition by the corresponding CN condition. (Here, in altering 23 and 29, (B_1) can be omitted.) For URCRN transformations, replace in the above set each URC condition by the corresponding URCRN condition. (Here, in altering 18, 19, and 20, (E_4^*) can be left unchanged; and in altering 23 and 29, (B_1) can be omitted.)

RC transformations.

Since (C_1) $\supset$ (A_1) and (C_2) $\supset$ (A_2), (A_1) and (A_2) are omitted.

30. NS RCRN → RC are (C_1), (D_1), and (F_1).
31. NS (RCN, RCURN) → RC are (C_1), (D_1), (D_2), (F_1), and (F_2).
32. NS RC → RC are (C_1), (D_1), (D_2), (D_3), (F_1), (F_2), and (F_3).
33. NS (BCN, BURCN, BURCRN) → RC are (C_1), (D_4), and (F_4).
34. NS (BC, BURC) → RC are (C_1), (D_3), (D_4), (F_3), and (F_4).
35. NS B → RC are (C_1), (D_5), and (F_5).
36. NS (CN, URCN, URCRN) → RC are (C_1), (C_2), (D_1), and (F_1).
37. NS (C, URC) → RC are (C_1), (C_2), (D_1), (D_3), (F_1), and (F_3).
38. NS UB → RC are (C_1), (C_2), (D_5), and (F_5).

For RCN transformations, replace in the above set each C condition by the corresponding CN condition. For RCRN transformations, replace in the above

set each RC condition by the corresponding RCRN condition. (Here, in altering 38, (C_1) can be replaced by (A_1).)

RCURN transformations. (See remark at head of preceding list.)

39. NS RCRN → RCURN are (C_1), (D_1), ($\bar{E}_1$), and (F_1).

40. NS RCURN → RCURN are (C_1), (D_1), (D_2), ($\bar{E}_1$), ($\bar{E}_2^*$), (F_1), and (F_2).

41. NS RCN → RCURN are (C_1), (D_1), (D_2), ($\bar{E}_1$), ($\bar{E}_2$), (F_1), and (F_2).

42. NS RC → RCURN are (C_1), (D_1), (D_2), (D_3), ($\bar{E}_1$), ($\bar{E}_2$), ($\bar{E}_3$), (F_1), (F_2), and (F_3).

43. NS (BCN, BURCN, BURCRN) → RCURN are (C_1), (D_4), ($\bar{E}_1$), and (F_4).

44. NS (BC, BURC) → RCURN are (C_1), (D_3), (D_4), ($\bar{E}_1$), ($\bar{E}_3$), (F_3), and (F_4).

45. NS B → RCURN are (C_1), (D_5), ($\bar{E}_1$), and (F_5).

46. NS (CN, URCN, URCRN) → RCURN are (C_1), (C_2), (D_1), ($\bar{E}_1$), and (F_1).

47. NS (C, URC) → RCURN are (C_1), (C_2), (D_1), (D_3), ($\bar{E}_1$), ($\bar{E}_3$), (F_1), and (F_3).

48. NS UB → RCURN are (C_1), (C_2), (D_5), ($\bar{E}_1$), and (F_5).

B transformations.

Since ($\bar{C}_1$) ⊃ (A_1) and ($\bar{C}_2$) ⊃ (A_2), (A_1) and (A_2) are omitted.

49. NS (B, BC, BCN, BURC, BURCN, BURCRN, RC, RCN, RCURN, RCRN) → B is ($\bar{C}_1$).

50. NS (UB, C, CN, URC, URCN, URCRN) → B are ($\bar{C}_1$) and ($\bar{C}_2$).

It is now possible to obtain NS $U \to V$ for each U of the 16 classes of sequences and each V of the set of classes BC, BCN, BURC, BURCN, BURCRN, BRC, BRCN, BRCURN, and BRCRN. To do this, simply add to the set of conditions for the corresponding transformation which does not involve boundedness, the condition or conditions from 49 or 50, as the nature of U requires.

Since ($\bar{C}_1$) ⊃ (C_1) ⊃ (B_1), ($\bar{C}_2$) ⊃ (C_2) ⊃ (B_2), ($\bar{C}_1$) ⊃ (A_1), and ($\bar{C}_2$) ⊃ (A_2), this process reduces merely to replacing, whenever they occur, (A_1) and (B_1) together, or (C_1), by ($\bar{C}_1$); and (A_2) and (B_2) together, or (C_2), by ($\bar{C}_2$). Certain of the resulting sets of conditions can be somewhat condensed. Thus (pp. 55–57 of $H_{1,2}$):

NS (CN, URCN, URCRN) → BURC are ($\bar{C}_1$), ($\bar{C}_2$), (D_1), and (E_1).
NS (C, URC) → BURC are ($\bar{C}_1$), ($\bar{C}_2$), (D_1), (D_3), (E_1), and (E_3).
NS (CN, URCN, URCRN) → BURCN are ($\bar{C}_1$), ($\bar{C}_2$), ($\bar{D}_1$), and (E_1).
NS (C, URC) → BURCN are ($\bar{C}_1$), ($\bar{C}_2$), ($\bar{D}_1$), ($\bar{D}_3$), (E_1), and (E_3).
NS (CN, URCN, URCRN) → BURCRN are ($\bar{C}_1$), ($\bar{C}_2$), (D_1), and ($\bar{E}_1$).
NS (C, URC) → BURCRN are ($\bar{C}_1$), ($\bar{C}_2$), (D_1), (D_3), ($\bar{E}_1$), and ($\bar{E}_3$).

The forms of the principal limits and of the row limits in all cases in which these exist can be read out of $H_{1,2}$ (e.g., p. 47, (7.1), etc.), if r^1 is therein replaced by x^1 wherever r^1 appears, and functional, rather than subscript, notation is used. (Thus, replace a_{r^1k} by $a_k(x^1)$.) This conclusion can be reached through the following considerations. If $F(x)$ is c, let $\mathfrak{S} \in \mathfrak{R}_1$. Then $\sigma = \lim_{m\to\infty} F(x:m)$.

And if x^1 is such that $\lim\limits_{g^2(x^2)\to\infty} F(x)$ exists, let $\mathfrak{S}^2 \epsilon \mathfrak{K}_1^2$ and let $x^1 = x^1\!:\!r^1$ for some $\mathfrak{S}^1 \epsilon \mathfrak{K}_1^1$ and some r^1. Then $\mathfrak{S} \epsilon \mathfrak{K}_1$, and, with $(m^1 = r^1)$, $\sigma(x^1) = \lim\limits_{m^2\to\infty} f(x:(m)) = \sigma_{r^1}$ for this $\mathfrak{S}$.

6. **Regularity conditions.** Whenever conditions on $\| f_k \|$ NS $U \to V$, where U and V involve convergence, are such that $\sigma = s$, this transformation will be called *regular*. Regularity conditions have been obtained above, incidentally, for each U which involves null convergence. The classes of type U to be yet considered are, then, RC, BURC, BC, URC, and C; those of class V are C, URC, RC, BC, BURC, and BRC. There have been derived above the 30 sets of conditions:

(1) NS RC $\to$ C, NS BURC $\to$ C, $\cdots$;

NS RC $\to$ URC, NS BURC $\to$ URC, $\cdots$; etc.

A condition necessary for regularity is that the limit be preserved for null sequences. Hence to the conditions (1) must be added, respectively, NS RCN $\to$ CN, NS BURCN $\to$ CN, $\cdots$; NS RCN $\to$ URCN, $\cdots$; etc.

Now by definition $(\bar{\mathrm{D}}_1) \supset (\mathrm{D}_1)$, $(\bar{\mathrm{D}}_2) \supset (\mathrm{D}_2)$, and $(\bar{\mathrm{D}}_4) \supset (\mathrm{D}_4)$. Examination of the relevant conditions now shows that[8] NS RC $\to$ C reg null, NS BURC $\to$ C reg null, $\cdots$; NS RC $\to$ URC reg null, $\cdots$; etc., are simply the sets of conditions (1), respectively, with (D_1), (D_2), and (D_4), whenever they occur, replaced by $(\bar{\mathrm{D}}_1)$, $(\bar{\mathrm{D}}_2)$, and $(\bar{\mathrm{D}}_4)$, respectively. Under the so-revised sets of conditions, the form of the principal limit of $F(x)$ is seen in each case to be $\sigma = L\cdot s$. Hence, finally, if there is introduced the

Regularity condition.

$$(\mathrm{D}_3^*) \qquad \lim_{g(x)\to\infty} \sum_{k=1}^{\infty} f_k(x) = 1,$$

then[9] NS RC $\to$ C reg, NS BURC $\to$ C reg, $\cdots$; NS RC $\to$ URC reg, $\cdots$; etc., are the sets of conditions (1), respectively, with (D_1), (D_2), (D_3), and (D_4), whenever they occur, replaced by $(\bar{\mathrm{D}}_1)$, $(\bar{\mathrm{D}}_2)$, (D_3^*), and $(\bar{\mathrm{D}}_4)$, respectively.

7. **Applications.** It will have been noticed that the sets $\mathfrak{E}^{(\nu)}$ and the set-sequences $\{\mathfrak{E}_j^{(\nu)}\}$ for which the theorems of this paper hold are very general. Upon the nature of these things, and upon them alone, depend the definitions of the various classes of functionals, UB, C, etc., and the conclusions expressed by the several theorems. Various specializations suggest themselves in obvious ways, among them, that, for $\nu = 1, 2, \cdots, l$, $\mathfrak{E}^{(\nu)}$ be the points of a metric space and $g^{(\nu)}$ a distance function defined with respect to some fixed point.

[8] *Reg null* means *with preservation of the limit for null sequences.*

[9] *Reg* means *regularly*, that is, so that the transformation be regular.

However, a set $\mathfrak{E}^{(\nu)} \equiv \mathfrak{E}^{(\nu+1)}$ ($\nu = 1, 2, \cdots, l-1$) for which metric ideas are non-essential (see the end of §2) reduces the present definitions and theorems to those of $H_{1,2}$; and applications to Euclidean spaces will be sufficiently suggestive of general metric applications. Definitions of $\mathfrak{E}^{(\nu)}$, $\mathfrak{E}_j^{(\nu)}$, and $g^{(\nu)}$ appropriate to a few typical applications will now be given.

I. *Transformations from sequence to sequence.* Such definitions have already been made at the end of §2. Attention may be called to the facts that under this specialization (C_1) and (C_2) reduce to ($\bar{C}_1$) and ($\bar{C}_2$), respectively; and BRC, BRCN, BRCURN, and BRCRN identify themselves, respectively, with RC, RCN, RCURN, and RCRN.

II. *Transformations from sequence to function over a Euclidean space.* Here, for $\nu = 1, 2, \cdots, l$, $\mathfrak{E}^{(\nu)}$ can be taken to be the points of a Euclidean space of q_ν dimensions, $\mathfrak{E}_j^{(\nu)}$ the subsets for which $|x^{(\nu)}| \geqq j-1$ ($j = 1, 2, \cdots$), and $g^{(\nu)}(x^{(\nu)}) \equiv x^{(\nu)}$ for each $x^{(\nu)}$.

What perhaps seems to be the most direct extension of the sequence-to-sequence transformation is provided by the following specialization. Let, for $\nu = 1, 2, \cdots, l$, $\mathfrak{E}^{(\nu)}$ be the set of non-negative real numbers, $\mathfrak{E}_j^{(\nu)}$ the subsets for which $x^{(\nu)} \geqq j-1$ ($j = 1, 2, \cdots$), and $g^{(\nu)}(x^{(\nu)}) \equiv x^{(\nu)}$ for each $x^{(\nu)}$.

If it is desired to treat of the behavior of $F(x)$ at a point P of some Euclidean space, a situation of somewhat the following sort is suggested. Let, for $\nu = 1, 2, \cdots, l$, $\mathfrak{E}^{(\nu)}$ be the points of a Euclidean space of q_ν dimensions, excluding some one point, $x_0^{(\nu)}$; $\mathfrak{E}_j^{(\nu)}$ the subsets for which $|x^{(\nu)} - x_0^{(\nu)}| < 1/j$ ($j = 2, 3, \cdots$); and $g^{(\nu)}(x^{(\nu)}) \equiv 1/|x^{(\nu)} - x_0^{(\nu)}|$.

Since the sets $\mathfrak{E}^{(\nu)}$ are mutually independent, the number of specializations which might be of interest is large.

Finally, consider in detail a very particular specialization for two transformations. Let it be required to state NS $F(t) \equiv \sum_{i,j=1}^{\infty} f_{ij}(t)s_{ij}$ exist for each t in $(0, \infty)$ and converge, as t tends to infinity, to $\lim_{i,j\to\infty} s_{ij}$, whenever this limit exists, (i) provided that $\{s_{ij}\}$ is bounded and (ii) in general. The transformations involved are (i) BC $\to$ C reg and (ii) C $\to$ C reg. Here $l = 1$, $\mathfrak{E}^{(1)}$ is the set of non-negative real numbers $x^{(1)} \equiv t$, $\mathfrak{E}_j^{(1)}$ can be taken to be the sets for which $t \geqq j-1$ ($j = 1, 2, \cdots$), and $g^{(1)}(t)$ can be defined to be equal to t for each t. Hence the relevant theorems become:

(i) *NS BC $\to$ C reg are*:

$$(A_1) \qquad \sum_{i,j=1}^{\infty} |f_{ij}(t)| < \infty \qquad (0 \leqq t < \infty);$$

$$(B_1) \qquad \sum_{i,j=1}^{\infty} |f_{ij}(t)| < A \textit{ for } t > B;$$

$$(D_3^*) \qquad \lim_{t\to\infty} \sum_{i,j=1}^{\infty} f_{ij}(t) = 1;$$

$$(\bar{D}_4)\qquad \lim_{t\to\infty}\sum_{i=1}^{\infty}|f_{ij}(t)| = 0 \qquad (j = 1, 2, \cdots),$$

$$\lim_{t\to\infty}\sum_{j=1}^{\infty}|f_{ij}(t)| = 0 \qquad (i = 1, 2, \cdots).$$

(ii) *NS C → C reg are*:

$$(A_1)\qquad \sum_{i,j=1}^{\infty}|f_{ij}(t)| < \infty \qquad (0 \leqq t < \infty);$$

$$(A_2)\qquad f_{ij}(t) = 0 \textit{ for } i > C_j(t) \qquad (0 \leqq t < \infty; j = 1, 2, \cdots),$$

$$f_{ij}(t) = 0 \textit{ for } j > C_i(t) \qquad (0 \leqq t < \infty; i = 1, 2, \cdots);$$

$$(B_1)\qquad \sum_{i,j=1}^{\infty}|f_{ij}(t)| < A \textit{ for } t > B;$$

$$(B_2)\qquad f_{ij}(t) = 0 \textit{ for } t, i > C_j \qquad (j = 1, 2, \cdots),$$

$$f_{ij}(t) = 0 \textit{ for } t, j > C_i \qquad (i = 1, 2, \cdots);$$

$$(\bar{D}_1)\qquad \lim_{t\to\infty} f_{ij}(t) = 0 \qquad (i, j = 1, 2, \cdots);$$

$$(D_3^*)\qquad \lim_{t\to\infty}\sum_{i,j=1}^{\infty} f_{ij}(t) = 1.$$

POMONA COLLEGE.

SUMS OF n-TH POWERS IN FIELDS OF PRIME CHARACTERISTIC

By Leonard Tornheim

A consequence of Waring's theorem on the representation of integers as sums of n-th powers is that every positive rational number is expressible as a finite number of n-th powers, the number required being less than a constant depending upon n. In the present paper we obtain similar theorems for fields of prime characteristic. The results tell which quantities are expressible as sums of n-th powers and how many n-th powers are needed.

Let F be a field of characteristic p, G the multiplicative group of non-zero elements, H the subgroup of all n-th powers, L the set of all elements expressible as sums of n-th powers, and K the set of all non-zero elements of L. We first prove

Lemma 1. *The set L is a field.*

Evidently L is closed under addition and multiplication. If $x = \sum_{i=1}^{r} x_i^n \neq 0$, then $-x = x + \cdots + x = (p - 1)x$ and $1/x = [(1/x)^n]x^{n-1}$. Thus L is a subfield of F and K is a subgroup of G containing H.

An example of a field for which $K \neq G$ is the finite field of four elements and $n = 3$. Here K has only the element 1.

Theorem 1. *Let F be a finite field. Every quantity in F which is expressible as a sum of n-th powers is a sum of n n-th powers.*

Let K_r be the set of all non-zero elements that are sums of r n-th powers; e.g., $H = K_1$. There exists a first subscript t for which $K_t = K$, $K_{t-1} < K$. Let $x = \sum_{i=1}^{t} x_i^n$ be in K_t, but not in K_{t-1}. Then $x' = \sum_{i=1}^{t-1} x_i^n$ is in K_{t-1} but not in K_{t-2}; otherwise $x = x' + x_t^n$ would be in K_{t-1}. Hence $K_{t-2} < K_{t-1}$. The argument is repeated for each $x^{(s)} = \sum_{i=1}^{t-s} x_i^n$ $(s = 2, \cdots, t - 1)$ to prove that $K_1 < K_2 < \cdots < K_t$.

If an element $y = \sum_{i=1}^{r} y_i^n$ is in K_r, the coset determined by y is in K_r, since every element of the coset has the form $z^n y = \sum_{i=1}^{r} (zy_i)^n$. It follows that each K_r, containing an element not in K_{r-1}, contains a coset not in K_{r-1}. Hence $t \leqq d$, where d is the index of H in K. Since $K = K_t$, every element expressible as a sum of n-th powers is a sum of $t \leqq d$ n-th powers.

It remains to prove that $d \leqq n$. Denote the index of H in G by m; then $d \leqq m$, since $K \leqq G$. The index m is equal to the number of distinct quantities

Received December 13, 1937.

of G whose n-th powers are equal, for example, to 1. If x is such a quantity, $x^n - 1 = 0$. The number of such quantities is equal to the number of roots of this equation of degree n. Hence $n \geqq m \geqq d$.

The prime field P of a field F is the field generated by the unity element 1. When F has characteristic p, P consists of $0, 1, \cdots, p-1$.

THEOREM 2. *Every quantity in an arbitrary field F of characteristic $p > n$ is a sum of $n(n+1)$ n-th powers.*

Consider the system

$$(a+r)^n = a^n + \binom{n}{1} a^{n-1} r + \cdots + r^n \qquad (r = 1, \cdots, n-1)$$

as a set of $n - 1$ linear equations in the $n - 1$ unknowns $a^{n-1}, a^{n-2}, \cdots, a$. The determinant of the coefficients is $D = \binom{n}{1}\binom{n}{2} \cdots \binom{n}{n-1} D'$, where

$$D' = \begin{vmatrix} 1 & 1^2 & \cdots & 1^{n-1} \\ 2 & 2^2 & \cdots & 2^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ n-1 & (n-1)^2 & \cdots & (n-1)^{n-1} \end{vmatrix},$$

and D' and the binomial coefficients are, except for sign, products and quotients of non-zero integers less than or equal to $n < p$. Thus D is in the prime field P and $D \neq 0$. Applying Cramer's method to solve the linear system for a, we get $a = D_1/D$, where D_1 is a linear homogeneous polynomial in the $n + 1$ quantities $a^n, 1^n, (a+r)^n$ $(r = 1, \cdots, n-1)$ with coefficients in P. Now each quantity in P is, by Theorem 1, a sum of at most n n-th powers. Hence a is a sum of $n(n+1)$ n-th powers.

The argument used for Theorem 2 can be modified to give a proof of

THEOREM 3. *Every quantity in an arbitrary field of infinite characteristic is a sum of n-th powers for n odd and also, if -1 is a sum of n-th powers, for n even. The number required is at most $g_n u_n(n+1)$, where g_n is the number of n-th powers needed to express any positive integer (the Waring constant) and u_n is the number of n-th powers in the expression of -1.*

We next prove

THEOREM 4. *Let $n = p^f - 1$. Every element in an arbitrary field F of characteristic p with more than $n + 1$ elements is a sum of n-th powers.*

The quantity $w = a^n - (a+1)^n$, where a is an arbitrary quantity of F, is in L. If $w \neq 0$, $a = (a^n - 1)/w - 1$ is in L. If $w = 0$, then $a = 1/(y-1)$, where $y \neq 1$ is a root of $x^n = 1$. Hence there are at most $n - 1$ quantities in G that are not in K.

If F is finite, K has $p^e - 1$ elements and G has $p^{ce} - 1$. By the preceding paragraph, $(p^{ce} - 1) - (p^e - 1) \leqq n - 1 = p^f - 2$, or $p^{ce} \leqq p^e + p^f - 2$. If $f \leqq e$, then $p^{ce} < 2p^e$, $c = 1$, and $K = G$. If $f > e$, then $p^{ce} < 2p^f - 2$,

$ce \leqq f$, $p^{ce} \leqq p^f = n + 1$. But F has more than $n + 1$ elements. Consequently the theorem is true for F finite.

If F is infinite, then K is infinite, the elements not in K being at most $n - 1$ in number. An element not in K determines a coset K' of G/K consisting of elements not in K equal in number to those in K. Since there are only a finite number of quantities in G not in K, no such coset K' exists. Hence $K = G$.

COROLLARY 1. *Every element in an infinite field of characteristic p is a sum of n-th powers if p does not divide n.*

In the theorem take $n_1 = p^m - 1$, where $m = \phi(n)$, ϕ being the Euler ϕ-function. Then n divides n_1. Every element, being a sum of n_1-th powers, is surely a sum of n-th powers.

In a perfect field of characteristic p every element has a p-th root.

COROLLARY 2. *There is only a finite number of perfect fields of finite characteristic for which not every element is a sum of n-th powers.*

Let $n = p^e q$, where p is a prime not dividing q. Let $n_1 = p^m - 1$, where $m = \phi(q)$; thus q divides n_1 and $n_1 \leqq n^n - 1$. If F, of characteristic p, has more than $n_1 + 1$ elements, every element of F is a sum of q-th powers, and since every element of F is a p-th power, every element is a sum of n-th powers. Hence every perfect field of characteristic p with more than $n_1 + 1$ elements is not an exceptional case, and consequently the same is true of any field of finite characteristic with more than n^n elements. But there is only a finite number of fields with less than n^n elements.

When n is a prime, there is at most one exceptional field.

THEOREM 5. *If q is a prime, there is at most one perfect field of characteristic p for which not every element is a sum of q-th powers. Such a field exists if and only if $q = 1 + p^f + p^{2f} + \cdots + p^{cf}$.*

For $p = q$, every element is a q-th power. For $p \neq q$, by Corollary 1 we may restrict our attention to finite fields.

For an exceptional case H is of prime index q in G. Since $K \neq G$, $K = H$. If the number of elements in G is $p^e - 1$, then q must divide $p^e - 1$. The set L has p^f elements, $f = e/k$, k an integer; K has $p^f - 1$ elements. Thus $q = \dfrac{p^{fk} - 1}{p^f - 1} = 1 + p^f + p^{2f} + \cdots + p^{(k-1)f}$. If another exception existed, q would also equal $\dfrac{p^{gs} - 1}{p^g - 1}$. Then $p^g(p^{kf} + p^{(s-1)g} - 1) = p^f(p^{gs} + p^{(k-1)f} - 1)$, and since $k > 1$, $s > 1$, we have $f = g$ and $s = k$.

Conversely, if $q = \dfrac{p^{fk} - 1}{p^f - 1}$, then the finite (Galois) field $F = GF(p^{fk})$ contains $L' = GF(p^f)$. The non-zero quantities of F form a multiplicative cyclic group G containing as subgroup the set K' of the non-zero quantities of L'. But H is a subgroup of G with index q in G since $x^q - 1$ has the q distinct solutions $1, z^r, z^{2r}, \cdots, z^{(q-1)r}$, where $r = p^f - 1$ and z is a generator of G. Now H has the same order as K' and, being subgroups of a cyclic group, they are iden-

tical, $H = K'$. Consequently the smallest field L containing H is L', and $K = K' = H$.

By using Theorem 5, the finite fields for which all elements are not sums of q-th powers, q being a prime, can be determined. For $q < 100$ they are given below.

q	3	5	7	13	17	31	73
GF	2^2	2^4	2^3	3^3	2^8	2^5, 5^3	2^9

UNIVERSITY OF CHICAGO.

COMPACT ABELIAN TRANSFORMATION GROUPS

By Deane Montgomery and Leo Zippin

1. **Introduction.** A topological group G is said to be a transformation group of a space R if to each element g of G there is associated a homeomorphism $g(R)$ taking R into itself in such a way that

(1) the function $g(x)$, g in G, x in R, is continuous simultaneously in g and x;

(2) if $g = g_1g_2$, then $g_1[g_2(x)] = g(x)$ for all x in R.

It follows from these conditions that to the identity element g_0 of G must correspond the identity homeomorphism

$$g_0(x) = x \qquad \text{for all } x \text{ in } R.$$

The set of elements g^* such that

$$g^*(x) = x \qquad \text{for all } x \text{ in } R$$

will be a closed invariant subgroup G^* and the factor group $\bar{G} = G/G^*$ may be defined in a very natural way as a transformation group[1] of R. This new transformation group is such that if $\bar{g}$ is not the identity, there is some point x for which $\bar{g}(x) \neq x$. A transformation group with this property we shall call an effective transformation group. We have thus seen that with any transformation group there is an associated effective transformation group very closely related to the original group.

A theorem, which is special but which is certainly the most complete of the present paper, follows:

An Abelian connected compact effective transformation group of Euclidean 3*-space* E *is isomorphic to the rotation group of a circle.*

From an earlier paper of ours[2] it follows that the coördinates in E can be chosen so that G becomes the familiar axial rotation group. Our analysis of the non-commutative connected transformation groups of 3-space being incomplete (we have disposed of the non-Abelian generalization of this theorem and will publish it shortly), we shall confine ourselves in this paper to a consideration of compact Abelian groups, and henceforth use the additive notation. When necessary, the group will be assumed to carry an invariant metric.[3]

The first part of this paper is concerned with some general theorems, only in

Received December 27, 1937.

[1] If $g \to \bar{g}$, then $\bar{g}(x) = g(x)$.

[2] *Periodic one-parameter groups in three-space*, Transactions of Am. Math. Soc., vol. 40(1936), p. 24.

[3] See D. van Dantzig, *Zur topologischen Algebra*, Math. Annalen, vol. 107(1933), especially pp. 615-616.

part preparatory to our final theorem, where the space R is understood to be a metric space and in some cases a Euclidean or locally Euclidean n-space.

2. Preliminary lemmas.

LEMMA 1. *Let the group H be the image of the compact Abelian group G under a continuous homomorphism by means of which the component of zero in G goes into the zero of H. Then H must be zero-dimensional.*

If the component of zero in G is the kernel of the homomorphism, the lemma is known to be true.[4] This, together with the fact that a factor group of a zero-dimensional group is zero-dimensional, is sufficient to complete the proof.

LEMMA 2. *Let the connected group H be the image of G under a continuous homomorphism f. Then the component of zero in G goes into all of H.*

Let G^* be the component of zero in G. The difference group $H - f(G^*)$ is connected. Now G can be mapped into $H - f(G^*)$ by a continuous homomorphism under which G^* goes into zero. Hence by Lemma 1, $H - f(G^*)$ is zero-dimensional. It must therefore be the identity group, and we conclude that $f(G^*) = H$.

Let G_i $(i = 1, \cdots, n)$ be a finite collection of closed subgroups of a compact group G. Let G_i^1 be the smallest closed subgroup including the subgroups $G_1, \cdots, G_{i-1}, G_{i+1}, \cdots, G_n$. The original set of subgroups is said to be *independent* if no G_i^1 includes G_i.

Remark. We note that the smallest subgroup containing a certain set of subgroups goes by a homomorphism of the whole group into the smallest subgroup containing the images of the subgroups of the set.

LEMMA 3. *Let $G_1, \cdots, G_n$ be closed subgroups of a compact group G and let f be a continuous homomorphism taking G into F in such a way that $f(G_1), \cdots, f(G_n)$ are independent in F. Then there is an e such that if h is a continuous homomorphism of G into a group H and G^h is in $S(G^f, e)$,[5] then $h(G_1), \cdots, h(G_n)$ are independent.*

We assume that G has an invariant metric. Since $f(G_1), \cdots, f(G_n)$ are independent, the groups $G_1, \cdots, G_n$ are independent. Let $H_i(G^f)$ be the set of elements of G which are in some coset (with respect to G^f) intersecting G_i^1. There is a positive number r such that every G_i contains a point whose distance from $H_i(G^f)$ is at least r. Now if $e = \frac{1}{2}r$, then $H_i(G^h)$ is in $S[H_i(G^f), \frac{1}{2}r]$ and this choice will serve in the lemma.

An interesting special case of the above may be obtained by letting f be the identity transformation.

LEMMA 4. *If a toral group has at least n connected closed independent subgroups, the group is at least n-dimensional.*

The proof is left for the reader, a toral group being, by definition, the direct product of a finite number of circle groups.

[4] See, for example, van Kampen, *Almost periodic functions and compact groups*, Annals of Math., (2), vol. 37(1936), pp. 78–91.

[5] This denotes all elements of G whose distance from G^f is less than e. The sets G^f and G^h are the kernels of the homomorphisms f and g.

LEMMA 5. *A compact connected group G is at least n-dimensional if and only if it contains n independent compact connected subgroups.*

The necessity will be shown first. Assume that G is a compact group which is at least n-dimensional. Then G contains a compact subgroup G^* such that $G - G^*$ is a toral group containing at least n circular summands. These are independent subgroups of $G - G^*$. Let f be the continuous homomorphism taking G into $G - G^*$. For each i let $F_i = f^{-1}(K_i)$. By a previous remark the compact groups F_i ($i = 1, \cdots, n$) are independent in G. Let G_i be the component of the identity of F_i. Then by Lemma 2 the groups G_i ($i = 1, \cdots, n$) are independent.

We turn now to the sufficiency of the condition. Let G be a group and G_i ($i = 1, \cdots, n$) the set of independent compact connected subgroups. The group G contains a subgroup G^* such that $G - G^*$ is toral. If G^* is chosen sufficiently small, the groups $G_1, \cdots, G_n$ go (by Lemma 3) into an independent set of connected subgroups of the torus. Hence the torus is at least n-dimensional and it follows that G must be at least n-dimensional.

3. Points on orbits of dimension at least k.

THEOREM 1. *Let G be a connected compact transformation group of a metric space R. Let K be all points of R whose orbits are at least k-dimensional. Then K is an open set.*

Let p be any point of K and let G_x for a variable point x in R denote the subgroup of G leaving the point x fixed. If a positive e is given, a d may be chosen so that, if $d(p, x) < d$, then $G_x \subset S(G_p, e)$. Since $G - G_p$ is at least k-dimensional, it contains k independent connected subgroups and these must be the images of certain connected independent groups $G_1, \cdots, G_k$ in G. Hence for an e satisfying the conditions of Lemma 3 and for the corresponding d, we see that x is in K provided $d(x, p) < d$.

4. Theorems on local connectedness.

THEOREM 2. *Let G be any compact group acting on a locally Euclidean n-space E. Let K be an $(n - 1)$-cell imbedded in E and let C be the component of $G(K)$ containing K. Then the set H of elements h such that $h(K) \subset C$ is an open and closed subgroup of G.*

The proof will be given for the case where E is Euclidean n-space. We verify first that H is a group. Suppose $h_1(K) \subset C$ and $h_2(K) \subset C$. Now $h_1(K)$ and K are in C. Therefore $(h_2 + h_1)(K)$ and $h_2(K)$ are in the same component of $G(K)$, and this component is C. It follows that $h_2 + h_1$ is in H. Furthermore, since $h_1(K)$ and K are in C, it is also true that K and $-h_1(K)$ are in the same component of $G(K)$, and thus $-h_1$ is in H.

H is closed because C is. We shall prove that H is open. Because H is a group, it will be sufficient to prove that the zero element of G is an inner point of H. Let K^* be the boundary of K. Let U be a neighborhood of zero in G so small that $U(K^*)$ does not separate $U(K)$ from infinity. Let a_1b_1 be an arc drawn from $E - U(K)$ to $U(K) - U(K^*)$ so that its only point in $U(K)$ is b_1.

There is an element g such that $\bar{g}(b_1)$ is in K. The arc $ab = \bar{g}(a_1b_1)$ is entirely in $E - U(K)$ except for its endpoint b, which is in K. Now for all sufficiently small g, $K + ab$ and $g(K)$ or K and $g(K + ab)$ must intersect. This is impossible unless K and $g(K)$ intersect. It is clear that very little modification in the proof is needed to reach the same conclusion for locally Euclidean spaces.

COROLLARY 1. *Let G be any compact group acting on E and let K be an $(n - 1)$-cell imbedded in E. Then $G(K)$ has a finite number of components.*

This is true because $G - H$ is a finite group.

THEOREM 3. *Let G be any connected compact group acting on a locally Euclidean n-space E and let K be an $(n - 1)$-cell in E. Then $G(K)$ has property*[6] *S, and is therefore locally connected.*

Let G^* be a small subgroup of G such that $G - G^*$ is a toral group. We may express K as the sum of a finite number of arbitrarily small $(n - 1)$-cells K_i so that by the preceding corollary $G^*(K_i)$ consists of a finite number of small components. Since $G - G^*$ may be divided into a finite number of small connected sets, the lemma follows.

COROLLARY 2. *If a connected compact group acts on a locally Euclidean n-space E, any $(n - 1)$-dimensional orbit is an $(n - 1)$-torus.*

This is true because any orbit is homomorphic to a compact group manifold and a connected locally connected compact $(n - 1)$-dimensional group is an $(n - 1)$-dimensional torus.

THEOREM 4. *If G is a zero-dimensional group acting on a connected locally Euclidean n-space E in such a way that every point has a finite orbit, there is an open and closed subgroup of G leaving every point of E fixed.*

Let $V_1, V_2, V_3, \cdots$ be a sequence of open and closed subgroups of G shrinking down to zero. Let E_n be the set of all points of E which are fixed under every element of V_n. By our hypothesis $E = E_1 + E_2 + \cdots$. Since the E_n's are closed, some one of them, say E_k, must contain an inner point. Let O be an open subset of E_k, and let g be any element of V_k. The transformation g which operates on E is pointwise periodic and is therefore periodic.[7] Since it leaves every point of O fixed, it leaves every point of the space fixed.[8] Because g is an arbitrary element of V_k, the proof is complete.

5. Non-existence of $(n - 1)$-dimensional orbits in n-space.

For any point a the set $G(a)$, which will sometimes be denoted by O_a, is called the orbit of a.

THEOREM 5. *If a connected compact k-dimensional group G acts on Euclidean n-space E $(n > 2)$, no orbit can be $(n - 1)$-dimensional.*

[6] A set is said to have property S if it can be expressed as the sum of a finite number of arbitrarily small closed and connected sets. See W. Sierpinski, *Sur une condition pour qu'un continu soit une courbe jordanienne*, Fundamenta Mathematicae, vol. 1(1920), pp. 44–60.

[7] D. Montgomery, *Pointwise periodic homeomorphisms*, American Journal of Mathematics, vol. 59(1937), p. 118.

[8] M. H. A. Newman, *A theorem on periodic transformations of spaces*, Quarterly Journal of Mathematics, (2), vol. 2(1931), p. 1.

Assume that an $(n - 1)$-dimensional orbit O_a exists. Let C be the component containing O_a in the set of points on $(n - 1)$-dimensional orbits. By Theorem 1, C is an open set. Let $\bar{G}$ be the difference group of G which operates effectively on C. Then $\bar{G}$ can contain no infinite zero-dimensional subgroups G^*. For the orbit $G^*(x)$, with x in C, can be regarded as a group which is a "subgroup" of the orbit $G(x)$ regarded as a group. Since $G(x)$ is toral, $G^*(x)$ is finite for all x of C, and hence by Theorem 4 G^* is finite.

It is known that a group which contains no infinite zero-dimensional group is a finite toral group.[9] We therefore have a toral group H of finite dimension k acting on a locally Euclidean space C.

Because O_a is $(n - 1)$-dimensional there must be subgroup H^* of H which leaves a fixed and whose dimension is $k - (n - 1)$. Let H^{**} be the component of the identity of H^*. The group H^{**} being a connected subgroup of a toral group is toral. Since the periodic elements of a toral group are everywhere dense, we may apply Newman's theorem to conclude that every point of C is fixed under H^{**}. Therefore we may define the action of the group $H - H^{**}$, which will be denoted by T^1, on C. Thus we may consider the set C as being acted on by the $(n - 1)$-dimensional toral group T^1. Let T^{1*} be the subgroup of T^1 leaving any point x of C fixed. The group T^{1*} must be zero-dimensional and therefore finite. Hence, since it leaves every point of O_x fixed, it must leave every point of C fixed. If any other element t of T^1 left any point of C fixed, it would also leave x fixed by the above argument. Hence any element not in T^{1*} moves every point of C. Now we define the action of $T = T^1 - T^{1*}$ on C in the customary manner and we have a toral group T acting on C in such a way that every element of T moves every point of C.

Let x and y be any two points of C such that y is outside O_x. (O_x, being an $(n - 1)$-dimensional torus, divides E into precisely two domains.) Let xy be an arc joining x and y and let this arc, except for the point x, be entirely outside O_x. A deformation of O_x into O_y may be defined in the following way. Let t be any point of xy and let p be any point of O_x. There is a unique g in T such that $g(x) = p$. The function $f(p, t) = g(t)$ defines a deformation of O_x into O_y, this deformation taking place in C and outside O_x. We see from this that O_y includes O_x.

We wish to show next that C must include all of E outside O_a. Suppose that this is not the case and let A be a point in the component of infinity of $E - C$. Let p be a point of O_a and let B be a point on the "outside" boundary of C. Since C is open, the orbit O_B must be less than $(n - 1)$-dimensional and hence can not separate E. There must exist an arc pA joining p and A which, except for p, is outside O_a and which does not touch O_B. Let $B_1, B_2, \cdots$ be a sequence of points outside O_a in C approaching B. By the preceding paragraph the arc pA must intersect O_{B_i} for each i. But since O_{B_i} approaches O_B,

[9] Pontrjagin, *The theory of topological commutative groups*, Annals of Mathematics, (2), vol. 35(1934), especially p. 386.

the arc must also intersect O_B. From this contradiction we conclude that C must include all points outside O_a.

We may now obtain a contradiction to our original assumption that O_a is an $(n - 1)$-dimensional orbit. We wish to show first that some cycle inside O_a does not bound inside O_a. By the duality theorem there exists a cycle Z_1 in $E - O_a$ which does not bound in $E - O_a$. If Z_1 is inside O_a, our task is accomplished. If Z_1 is outside O_a, it links a cycle Z_2 in O_a, and from our investigation of the way in which orbits may be deformed, we see that if y is a point of C inside O_a, then Z_2 may be deformed inside O_a to lie on O_y. In any case, therefore, there is a cycle Z inside O_a which does not bound inside O_a. The cycle Z must link a cycle Z^* in O_a. From the way in which O_a may be deformed it is clear that if y is any point outside O_a, then Z^* may be deformed outside O_a to a position in O_y. In this position Z^* still links Z. Hence as y approaches infinity, O_y must always intersect a fixed finite set bounded by Z. This is impossible, and the contradiction completes the proof of the theorem.

6. Zero-dimensional groups in two-dimensional manifolds.

In this section we shall show that if a zero-dimensional group G acts effectively on a 2-dimensional manifold, then G must be finite. In view of previous remarks it is clear that the theorem might be stated as follows: if G is a zero-dimensional group acting on a connected 2-dimensional manifold and if G^* is the subgroup of elements leaving every point of E fixed, then $G - G^*$ is finite, or, what amounts to the same thing, G^* is open. We sometimes use one formulation and sometimes the other.

LEMMA 6. *If a zero-dimensional group G acts effectively on a simple closed curve C, then G is finite.*

Let x be any point of C and y a point of C not in $G(x)$. Let $U(z)$ be an open arc of C containing x and no point of $G(y)$. We shall see that for every point z of $U(z)$, $U(z) \cdot G(z)$ consists of at most two points. For suppose that z_1, z_2, and z_3 are distinct points of some $U(z) \cdot G(z)$. Then one of these points, say z_2, is separated from every point of $G(y)$ by z_1 and z_3. Therefore any arc $z_2 y^1$ of C joining z_2 to a point of $G(y)$ meets z_1 or z_3.

Now if the orbits of points in C are taken as points of an auxiliary space C^*, there is an arc in C^* joining the point $G(z)$ and the point $G(y)$. Therefore, there is an arc in C which joins z_2 to some point of $G(y)$ and which has z_2 only in common with $G(z)$. This contradiction shows that $U(z) \cdot G(z)$ contains at most two points. It follows at once by an application of the Heine-Borel theorem that in every $G(z)$ there are at most $2N$ points, where N is the finite number of open sets $U(z)$ covering C. We now apply Theorem 4.[10]

THEOREM 6. *If a zero-dimensional group G acts on a connected 2-dimensional locally Euclidean space E, then G must contain an open subgroup leaving all points of E fixed.*

[10] This is a special case of a theorem (known to us when the abstract of this paper was prepared) on "continuous decompositions" of the line and circle into totally disconnected decomposition sets. See G. T. Whyburn, this Journal, vol. 3(1937), pp. 370–381.

Consider a domain D of E which is homomorphic to a plane and let D^1 be the boundary of D. Let C^1 be a simple closed curve in D and let C be the interior of C^1 (in D). Let x be any point in C. There exists an open and closed subgroup G^* of G such that $G^*(x)$ is in C. Let K be a continuum in C containing $G^*(x)$. We may suppose G^* so chosen that $G^*(C^1)$ belongs to D and contains no point of K.

The set $G^*(K)$ is a continuum having no point in common with $G^*(C^1)$ and belonging entirely to C. The component of $E - G^*(C^1)$ containing $G^*(K)$ is invariant under G^*. The boundary B^1 of this component is likewise invariant under G^* and separates x from points of D^1. The component of $E - B^1$ which contains D^1 is invariant under G^*. The boundary B^{11} of this component is invariant under G^* and separates x from D^1.

We see now that B^{11}, a subset of $G^*(C^1)$, is the common boundary of at least two domains. It is therefore a simple closed curve.[11] By Lemma 6 some open and closed subgroup G^{**} of G^* leaves every point of B^{11} fixed. It will now be shown that G^{**} leaves any point z of E fixed.

There must exist an arc azb such that a and b are on B^{11} and the arc azb separates that domain of $E - B^{11}$ in which it lies. In particular, in the closure of this domain it separates a pair of points c and d of B^{11}. By essentially the argument above $G^{**}(azb)$ contains an arc which is invariant under G^{**} and which joins a and b. The points a and b are fixed under G^{**} and it follows that every point of the invariant arc is fixed under G^{**}. We see that the invariant arc coincides with azb and hence that z is a fixed point.[12]

7. Connected groups in locally Euclidean 3-spaces.

THEOREM 7. *If G is a compact, connected, Abelian, effective transformation group of a 3-dimensional connected locally Euclidean space E, then G is toral.*

We begin with a brief outline of the method of proof. The group G must contain a one-parameter dense subgroup G'.[13] Suppose that G is not toral and let G^* be any closed infinite zero-dimensional subgroup.[14] Now under the action of the one-parameter subgroup G', E is filled with a regular family of curves in the sense of H. Whitney.[15] Let p be any point of E. We select an appropriate cross-section (Whitney, loc. cit.) at p of these orbits and show that

[11] R. L. Moore, *Foundations of Point Set Theory*, New York, 1932, p. 216.

[12] See footnote 10.

[13] We are not aware that this has ever been stated in the literature, although it must be fairly well known. It follows, for example, from the explicit "pseudo-basis" for a compact metric Abelian group given by J. W. Alexander, Annals of Mathematics, (2), vol. 35(1934), pp. 392 and 393. Thus the element $b = \sum \eta_i \beta_i$, where the coefficients η_i form a countable set (whether it is finite or infinite depending on the dimension of the group) of rationally independent reals (mod 1), generates a subgroup of this sort. This follows essentially from the Kronecker theorem. That every compact metric Abelian group possesses such a pseudo-basis is not shown by Alexander but follows, of course, from the work of Pontrjagin.

[14] In Alexander's formulation this is the group generated by the elements $b_1, b_2, \cdots$.

[15] *Regular families of curves*, Annals of Mathematics, (2), vol. 34(1933), p. 244.

a subgroup G^{**} of G^* can be construed as a transformation group of the section. We verify Whitney's statement[16] that the section is locally planar and by means of the dual rôles of G^{**} deduce the theorem.

We now present the proof in detail. Let p be any point of E and let G^* be as above. We wish to show that $G^*(p)$ is finite. If $G^*(p)$ is not finite, p is certainly moved by G^* and hence by G'. There exists an interval J of G', whose endpoints may be denoted by t' and $t'' = -t' > 0$, such that $J(p)$ is an arc $p'pp''$, $p' = t'(p)$, $p'' = t''(p)$. Whitney (loc. cit.) has shown that there exist a real-valued function (we shall call it $W(x)$) defined for all points x and a neighborhood S of the point p such that $W(x)$ is continuous in S and increases throughout S as we move along orbits of G' in the sense of increasing parameter. In particular, then

$$W(p') < W(p) < W(p''),$$

since by diminishing J, if necessary, we may arrange that $p'pp''$ is in S.

It is easily seen that there exist mutually exclusive connected neighborhoods $S(p')$, $S(p)$, $S(p'')$, of which the first and last may be supposed "spherical" and the second may be supposed *invariant* under an appropriate open subgroup G^{**} of G^*, such that

$$\text{(i)} \qquad \overline{S(p')} + J\{\overline{S(p)}\} + \overline{S(p'')} \subset S,$$

$$\text{(ii)} \qquad t'\{S(p)\} \subset S(p') \text{ and } t''\{S(p)\} \subset S(p''),$$

$$\text{(iii)} \qquad W(x) \lessgtr W(p) \text{ if } x \subset \begin{matrix} S(p') \\ S(p'') \end{matrix}.$$

The set $J\{S(p)\}$ is open and connected and it is invariant under G^{**}, since $g\{S(p)\}$ (for any g of G) must be invariant under G^{**}.

Let W be the set of points of $S(p)$ at which the function $W(x)$ has the value $W(p)$.[17] For any point x of $S(p)$ a subarc of $G'\{x\}$ which belongs to S contains at most one and, if it includes $J\{x\}$, at least one point of W. Moreover (from the continuity of $W(x)$ and from the preceding remark), the set $J\{W\}$ is the topological product of the set W (regarded as a space) and the linear interval.

Let g be an arbitrary element of G^{**}, w a point of W. Then $g\{J\{w\}\}$ is the arc $J\{g\{w\}\}$ and contains a unique point, which we shall designate by $g[w]$, of W:

$$g[w] = W \cdot g\{J\{w\}\} = W \cdot J\{g\{w\}\}.$$

By the correspondence $w \to g[w]$, the elements of G^{**} become single-valued transformations of W into itself.

The transformation $g[w]$ is continuous simultaneously in g and w. For, let $g_n \to g$, $w_n \to w$. With each t of J the sequence $g_n\{t\{w_n\}\}$ converges to the point $g\{t\{w\}\}$, which varies continuously with t. The totality of these points is precisely the set $g\{J\{w\}\}$ and this is the sequential limiting set of the

[16] Whitney informs us that he is about to publish a proof.

[17] This is, for the case in hand, the Whitney cross-section.

$g_n\{J\{w_n\}\}$. Now the points $g_n[w_n]$ have at least one limit point in $\overline{S(p)}$; any such point must belong to $g\{J\{w\}\}$ and the value of the function $W(x)$ at this point must be $W(p)$. There is just one point satisfying these conditions: $g[w]$.

We now prove that if $g = g_1 + g_2$ (elements of G^{**}), then

$$g[w] = g_1[g_2[w]],$$

where these transformations of W are compounded as ordinarily. We observe that the arcs $J\{g_2\{w\}\}$ and $J\{g_2[w]\}$ have in common the point $g_2[w]$. Therefore the arcs $g_1\{J\{g_2\{w\}\}\}$ and $g_1\{J\{g_2[w]\}\}$ have in common the point $g_1\{g_2[w]\}$. Their sum is therefore a connected set belonging to S, since this contains each arc, and to the orbit $G'\{w\}$. This sum can contain at most one point of W. On the other hand, each arc meets W in precisely one point so that they must each meet W in the same point, namely, $g_1[g_2[w]]$. But it is clear that the arc $g_1\{J\{g_2\{w\}\}\}$ is identically the arc $g\{J\{w\}\}$, and the proof is accomplished.

It is obvious that the transformation of W induced by the identity element of G^{**} is merely the identity: $g_0[w] = w$. Then it follows from the preceding section that to inverse elements of G^{**} correspond inverse transformations of W and from this that all the transformations $g[w]$ are auto-homeomorphisms of W, since these transformations have all been shown to be continuous as well as single-valued. By the simultaneously continuous association $g[w]$ of points of W and elements of G^{**}, the group G^{**} has been made a transformation group on the "space" W.

We shall now interrupt our argument to establish rapidly that a Whitney section is locally planar (in 3-space).[18] Essentially as in the foregoing discussion (but discarding the group G^*) let p be the point at which the section W is taken, S the neighborhood where the W-integral increases, J the appropriate set of parameter values t: $t' \leqq t \leqq t'' = -t'$ (> 0); $S(p)$ a "solid" spherical neighborhood of p such that it, $S'(p') = t'\{S(p)\}$, and $S''(p'') = t''\{S(p)\}$ are mutually exclusive. Now let W^* be the component of W in $S(p)$, containing p.

Let L denote an arc ab of W^* which (as a special case) may degenerate to the single point $a = b$. Then either $J(L)$ is, degenerately, an arc with endpoints $t'(a)$ and $t''(a)$ or it is a closed 2-cell bounded by the four edges: $J(a)$, $t''(L)$, $J(b)$, $t'(L)$. Now in either case if V denotes an arbitrary neighborhood of the point a (which we may take to lie inside $S(p)$), there exists a neighborhood U of this point such that if x and y are any two points of $U \cdot W^*$ not in $J(L)$ there exists an arc xy in V not meeting $J(L)$.[19]

If z is an arbitrary point of the arc xy, there is associated with it uniquely

[18] See Whitney, this Journal, vol. 4(1938), pp. 222–226. The proof here given, while not elementary, has the merit that it can easily be adapted to give considerable information about corresponding sections in spaces which possess locally the combinatorial properties of n-space.

[19] P. Alexandroff, *On local properties of closed sets*, Annals of Mathematics, (2), vol. 36 (1935), pp. 1–35.

and continuously a point z^* on W^*, and the set of these points contains an arc xy which lies in W^*. By a more restrictive choice of the neighborhood U (bringing the points x and y "closer" to a) this arc may be made to lie in a preassigned neighborhood of a in W^*. When L is a point it follows that the local complement in W^* of every point is arcwise connected (so that W^* is certainly locally connected, as shown by Whitney, loc. cit.) and that the same statement is also true for any arc L of W^*. Then, clearly, no arc of W^* can separate W^* (even locally).

Now let C denote a simple closed curve of W^*. Let T denote the closed set $S'(p') + J(C) + S''(p'')$. This can obviously be expressed as the sum of two closed sets T_1 and T_2 of which the first contains $S'(p')$ and the "negative" half of the cylinder $J(C)$ while the second contains $S''(p'')$ and the "positive" half of this cylinder. The two sets have precisely the set C in common. Recalling that $t'(C) \subset S'(p')$ and is homologous to zero there, since this is a solid sphere (similarly with $t''(C)$), it follows at once from the "Generalized Phragmén-Brouwer Theorem" of P. Alexandroff that T must separate E. It is not difficult to see that T must separate E between points of W^* near an arbitrary point of C. It is an immediate consequence of this that C must separate W^*. We know, finally, that W^* is homeomorphic to an open connected subset of the two-sphere.[20]

Returning to W and the group G^{**} operating on it, we see by Theorem 6 that there is a subgroup G^{***} such that for every element g of G^{***}, $g[w] = w$. This means, from the definition of $g[w]$, that the point $g\{w\}$ belongs to the arc $J\{w\}$. Hence the group G^{***} is mapped homomorphically into a subgroup of the interval J. But this interval contains no subgroups, so that G^{***} is mapped into the identity element; i.e., for each g, $g\{w\} = w$. Then, in particular, for the original point p and for every g of G^{***}, $g\{p\} = p$ and the theorem is proved.

Theorem 8. *If a compact connected Abelian group G acts effectively on Euclidean* 3*-space, then G is the circle group* (*and is the axial rotation group of* 3*-space*).

We already know that G is toral. If G is not a circle group, it must contain at least two distinct circle subgroups T_1 and T_2. We shall obtain a contradiction to this.

We know that under the action of T_1, say, the orbit space E^* of E is a closed half-plane, its edge F^* corresponding to the axis F of fixed points. Now because the groups T_1 and T_2 commute, the group T_2 can be shown to be a transformation group on the space E^* which leaves points of F^* fixed. Hence by Newman's theorem (since the elements of finite order are dense in T_2) the group T_2 leaves every point of E^* fixed. Reversing the rôles of T_1 and T_2, we see that for every point x of E, $T_1(x) = T_2(x)$, that is, the orbits under T_1 and T_2 coincide and the set F contains those points which are fixed under both T_1 and T_2. We note that every point of F is fixed under the group T^* generated by T_1 and T_2, it being clear that T^* is two-dimensional.

[20] L. Zippin, *On continuous curves and the Jordan curve theory*, American Journal of Mathematics, vol. 52(1930), pp. 331–350.

Let x be any moving point not in F. Let T_x denote the subgroup of T^* leaving x fixed. Because $T^*(x)$ is one-dimensional it follows that T_x is one-dimensional and that it must contain a circular subgroup C. We see that every point of F is fixed under C and that in addition the point x is fixed under C. This is impossible (Montgomery and Zippin, loc. cit.) unless every point of the space is fixed under C, i.e., unless T^* is not effective. Therefore the original group G could not have been effective.

SMITH COLLEGE AND NEW YORK UNIVERSITY.

CERTAIN INTEGRALS AND INFINITE SERIES INVOLVING ULTRASPHERICAL POLYNOMIALS AND BESSEL FUNCTIONS

By Hsien-yü Hsü

1. **Introduction.** Let $P_n^{(\lambda)}(x)$ denote the ultraspherical polynomials defined by the generating function $(1 - 2xw + w^2)^{-\lambda} = \sum_{n=0}^{\infty} P_n^{(\lambda)}(x)w^n$ [cf. **6**, p. 50; **7**, p. 329; **4**, p. 37].[1] The following considerations are devoted to the discussion, first, of the integral formula

$$
\begin{aligned}
D_1(\lambda; l, m, n) &\equiv \int_{-1}^{1} (1 - x^2)^{\lambda - \frac{1}{2}} P_l^{(\lambda)}(x) P_m^{(\lambda)}(x) P_n^{(\lambda)}(x)\, dx \\
&= \frac{2^{1-2\lambda}}{\{\Gamma(\lambda)\}^2} \frac{\pi}{s + \lambda} \frac{\Gamma(s + 2\lambda)}{\Gamma(s + 1)} \\
&\qquad \frac{\binom{s - l + \lambda - 1}{s - l}\binom{s - m + \lambda - 1}{s - m}\binom{s - n + \lambda - 1}{s - n}}{\binom{s + \lambda - 1}{s}} \quad \text{or } 0,
\end{aligned} \tag{1.1}
$$

second, of the infinite expansion

$$
\begin{aligned}
D_2(\lambda; \alpha, \beta, \gamma) &\equiv \sum_{n=0}^{\infty} (n + \lambda) \left\{\frac{\Gamma(n + 1)}{\Gamma(n + 2\lambda)}\right\}^2 P_n^{(\lambda)}(\cos \alpha) P_n^{(\lambda)}(\cos \beta) P_n^{(\lambda)}(\cos \gamma) \\
&= 2^{-2\lambda} \pi \{\Gamma(\lambda)\}^{-4} \{\sin \alpha \sin \beta \sin \gamma\}^{1-2\lambda} \\
&\qquad \left\{\sin \frac{\alpha + \beta + \gamma}{2} \sin \frac{\beta + \gamma - \alpha}{2} \sin \frac{\gamma + \alpha - \beta}{2} \sin \frac{\alpha + \beta - \gamma}{2}\right\}^{\lambda - 1} \\
&\qquad \text{or } 0,
\end{aligned} \tag{1.2}
$$

third, of the integral formula

$$
S(\nu; a, b, c) \equiv \int_0^{\infty} J_\nu(ax) J_\nu(bx) J_\nu(cx) x^{1-\nu}\, dx = \frac{2^{\nu-1} \Delta^{2\nu-1}}{\Gamma(\frac{1}{2})\Gamma(\nu + \frac{1}{2})(abc)^\nu} \quad \text{or } 0. \tag{1.3}
$$

In (1.1) we have $\lambda > -\frac{1}{2}$ and the numbers l, m, n are arbitrary non-negative integers. The first of the two given values holds if $l + m + n$ is even, $l + m + n = 2s$, and a triangle exists with the sides l, m, n; and the second value holds in every other case.[2]

In (1.2) we have $\lambda > 0$ and the parameters α, β, γ are arbitrary positive

Received January 12, 1938.

[1] The bold face numbers refer to the bibliography at the end.

[2] In case $\lambda = 0$ the formula needs a slight modification.

numbers less than π. The first, or the second, value holds according as a spherical triangle can, or cannot, be drawn with the sides α, β, γ.

In (1.3), $J_\nu(x)$ denotes Bessel's function of order ν, where $\nu > -\frac{1}{2}$, and the parameters a, b, c are positive numbers. The first, or the second, value holds according as a triangle can, or cannot, be drawn with the sides a, b, c. The area of this triangle is denoted by Δ.

Formula (1.1) remains true if the area of the triangle becomes 0. The series (1.2) and the integral (1.3) are in this case properly divergent for $\lambda \leqq 1$ and for $\nu \leqq \frac{1}{2}$, respectively; they are, however, convergent (and $= 0$) for $\lambda > 1$ and for $\nu > \frac{1}{2}$, respectively. In the subsequent considerations, except in the proof of (1.1), we assume that *the area of the corresponding triangle is greater than zero.*

Formula (1.1), in the case of Legendre polynomials, for which $\lambda = \frac{1}{2}$, was stated without proof by N. M. Ferrers. J. C. Adams proved it by mathematical induction. Later proofs for this special case have been given by I. Todhunter and F. Neumann. The general formula (1.1) has been pointed out by J. Dougall.[3] However, he did not give a proof for it. He merely stated that the method, by which he obtained this result, "is long and far from neat". [See the references in **1**, p. 47; furthermore, **5**.]

Formula (1.2) is also due to Dougall [**1**]; his proof is based on potential theory. Formula (1.3) is a classical theorem due to Sonine [see **6**, p. 411]. Dougall also bases a proof for Sonine's theorem on considerations analogous to those in the proof of (1.2). Furthermore, he remarks that both formulas (1.1) and (1.2) contain Sonine's result as a "limiting case"; however, he does not give any further explanation of this remark.

It is the purpose of this paper:

(a) to give a simple proof for (1.1), based on mathematical induction; this is the first proof to be given for this theorem, since Dougall's proof, as mentioned above, was never published;

(b) to give a simple proof for (1.2) entirely different from, and shorter than, that of Dougall;

(c) to show that from formula (1.1), by means of a limiting process, Sonine's formula (1.3) can be derived;

(d) similarly, to show that, as in (c), (1.3) can be derived from (1.2).

In the limiting process of (c) and (d), we use Mehler's formula and an asymptotic formula of Stieltjes (see below). Another arrangement of proof is possible by use of a formula of "Hilb's type" [**4**, p. 44]. However, the methods given in this paper are of a more elementary character. A proof of Sonine's theorem (1.3) can also be given by a procedure similar to that given for (1.2) [cf. (b)] by making use of Hankel's inversion formula.

2. **Proof of Dougall's integral formula** (1.1). It is easily seen from the symmetry property of the ultraspherical polynomials that the integral in

[3] Professor G. N. Watson kindly called my attention to this paper [**1**].

question vanishes for $l + m + n$ odd; it also vanishes on account of the orthogonality property, when $l + m + n = 2s$ and at the same time a triangle with the sides l, m, n cannot be constructed. In what follows we shall therefore assume that $l + m + n = 2s$, and the triangle condition (in the wider form) $l \leqq s$, $m \leqq s$, $n \leqq s$ is satisfied. We make use of the recurrence formula

$$P_n^{(\lambda)}(x) = a_n^{(\lambda)} x P_{n-1}^{(\lambda)}(x) - b_n^{(\lambda)} P_{n-2}^{(\lambda)}(x). \tag{2.1}$$

For sake of brevity, we write a_n, b_n, instead of $a_n^{(\lambda)}$, $b_n^{(\lambda)}$, respectively.

First, we will discuss the case $l = 0$. Then $m = n$ is the only case in which the triangle condition is satisfied. We then have [6, p. 367]

$$D_1(\lambda; 0, n, n) = \int_{-1}^{1} (1 - x^2)^{\lambda - \frac{1}{2}} \{P_n^{(\lambda)}(x)\}^2 \, dx = \frac{2^{1-2\lambda}}{\{\Gamma(\lambda)\}^2} \frac{\pi}{n + \lambda} \frac{\Gamma(n + 2\lambda)}{\Gamma(n + 1)}, \tag{2.2}$$

whence the statement (1.1) follows.

The general case can be established by mathematical induction. Whatever the values of m and n may be, the statement holds for $l = 0$; we assume it to hold for $l - 1$, and prove it for l.

By the recurrence formula, we have for $l \geqq 1$, $b_1 = 0$,

$$\begin{aligned} D_1(\lambda; l, m, n) &= \int_{-1}^{1} (1 - x^2)^{\lambda - \frac{1}{2}} [a_l x P_{l-1}^{(\lambda)}(x) - b_l P_{l-2}^{(\lambda)}(x)] P_m^{(\lambda)}(x) P_n^{(\lambda)}(x) \, dx \\ &= a_l \int_{-1}^{1} (1 - x^2)^{\lambda - \frac{1}{2}} P_{l-1}^{(\lambda)}(x) x P_m^{(\lambda)}(x) P_n^{(\lambda)}(x) \, dx \\ &\qquad - b_l D_1(\lambda; l - 2, m, n). \end{aligned} \tag{2.3}$$

In the first term of the last member we apply the recurrence formula again, and find

$$\begin{aligned} D_1(\lambda; l, m, n) &= a_l a_{m+1}^{-1} \int_{-1}^{1} (1 - x^2)^{\lambda - \frac{1}{2}} P_{l-1}^{(\lambda)}(x) \{P_{m+1}^{(\lambda)}(x) + b_{m+1} P_{m-1}^{(\lambda)}(x)\} P_n^{(\lambda)}(x) \, dx \\ &\qquad - b_l D_1(\lambda; l - 2, m, n) \\ &= a_l a_{m+1}^{-1} D_1(\lambda; l - 1, m + 1, n) + a_l a_{m+1}^{-1} b_{m+1} D_1(\lambda; l - 1, m - 1, n) \\ &\qquad - b_l D_1(\lambda; l - 2, m, n). \end{aligned} \tag{2.4}$$

To the terms of the right member of the last expression, we can apply the theorem. The result, apart from the common factor

$$\frac{2^{1-2\lambda}}{\{\Gamma(\lambda)\}^2} \frac{\pi}{s + \lambda} \frac{\Gamma(s + 2\lambda)}{\Gamma(s + 1)}$$

$$\frac{\binom{s - l + \lambda - 1}{s - l} \binom{s - m + \lambda - 1}{s - m} \binom{s - n + \lambda - 1}{s - n}}{\binom{s - \lambda + 1}{s}}, \tag{2.5}$$

where $2s = l + m + n$, is[4]

$$a_l a_{m+1}^{-1} \frac{s-l+\lambda}{s-l+1} \frac{s-m}{s-m+\lambda-1} + a_l a_{m+1}^{-1} b_{m+1} \frac{s+\lambda}{s+2\lambda-1} \frac{s-n}{s-n+\lambda-1}$$
$$- b_l \frac{s+\lambda}{s+2\lambda-1} \frac{s-n}{s-n+\lambda-1} \frac{s-l+\lambda}{s-l+1} \frac{s-m}{s-m+\lambda-1}.$$

We have in the case of the ultraspherical polynomials

$$a_l = \frac{2(l+\lambda-1)}{l}, \quad l \geqq 1; \qquad b_l = \frac{l+2\lambda-2}{l}, \quad l \geqq 2; \qquad b_1 = 0.$$

Thus, we find

$$\frac{m+1}{m+\lambda} \frac{l+\lambda-1}{l} \frac{s-l+\lambda}{s-l+1} \frac{s-m}{s-m+\lambda-1}$$
$$+ \frac{l+\lambda-1}{l} \frac{m+2\lambda-1}{m+\lambda} \frac{s+\lambda}{s+2\lambda-1} \frac{s-n}{s-n+\lambda-1} \tag{2.6}$$
$$- \frac{l+2\lambda-2}{l} \frac{s+\lambda}{s+2\lambda-1} \frac{s-n}{s-n+\lambda-1} \frac{s-l+\lambda}{s-l+1} \frac{s-m}{s-m+\lambda-1} = 1.$$

(In case $l = 1$, the third term of the left member must be replaced by 0. Furthermore, we have in this case by (2.4) either $n = m - 1$ or $n = m + 1$.) Indeed, this equation is satisfied by five values of λ: namely, $\lambda = -s, l - s, 1 - l, 1, \infty$. Here we consider l, m, n and s as arbitrary parameters. Hence, we have established the statement (1.1).

We notice that a formula of the type (2.4) holds in general, if $(1 - x^2)^{\lambda-\frac{1}{2}}$ in (1.1) is replaced by an arbitrary weight function $w(x)$, for which $w(-x) \equiv w(x)$, and $P_n^{(\lambda)}(x)$ by the corresponding orthogonal polynomials.

3. **Proof of Dougall's expansion formula** (1.2). Let us consider a spherical triangle with two fixed sides β and γ, $0 < \beta < \pi$, $0 < \gamma < \pi$; the angle A between β and γ should vary from 0 to π. Then the opposite side α can be calculated from

$$\cos \alpha = \cos \beta \cos \gamma + \sin \beta \sin \gamma \cos A. \tag{3.1}$$

If A increases from 0 to π, the expression of the right member changes from $\cos(\beta - \gamma)$ to $\cos(\beta + \gamma)$. We assume that $0 \leqq \alpha \leqq \pi$; then

(i) if $\beta + \gamma \leqq \pi$, α increases from $|\beta - \gamma|$ to $\beta + \gamma$;

(ii) if $\beta + \gamma > \pi$, α increases from $|\beta - \gamma|$ to $2\pi - (\beta + \gamma)$.

[4] In case $s = m$, the triangle condition is not satisfied in the first term; thus the corresponding D_1 has to be replaced by 0. Similarly, if $s = n$ in the second term, and $s = m$ or $s = n$ in the third term.

From the addition theorem for ultraspherical polynomials, we find by integration [**6**, p. 369]

$$\int_0^\pi P_n^{(\lambda)}(\cos\alpha)\sin^{2\lambda-1}A\,dA = 2^{2\lambda-1}\{\Gamma(\lambda)\}^2\frac{\Gamma(n+1)}{\Gamma(n+2\lambda)}P_n^{(\lambda)}(\cos\beta)P_n^{(\lambda)}(\cos\gamma). \tag{3.2}$$

Now we introduce α instead of A as the variable of the integration. Since

$$\begin{aligned}\sin\beta\sin\gamma\sin A &= \sin\beta\sin\gamma\left[1-\left(\frac{\cos\alpha-\cos\beta\cos\gamma}{\sin\beta\sin\gamma}\right)^2\right]^{\frac{1}{2}}\\ &= 2\left[\sin\frac{\alpha+\beta+\gamma}{2}\sin\frac{\beta+\gamma-\alpha}{2}\sin\frac{\gamma+\alpha-\beta}{2}\sin\frac{\alpha+\beta-\gamma}{2}\right]^{\frac{1}{2}} \equiv 2E^{\frac{1}{2}}\end{aligned} \tag{3.3}$$

and $\sin\beta\sin\gamma\sin A\,dA = \sin\alpha\,d\alpha$, we find for (3.2)

$$\begin{aligned}&\int_0^\pi P_n^{(\lambda)}(\cos\alpha)\sin^{2\lambda-2}A\sin A\,dA\\ &\qquad= 2^{2\lambda-2}(\sin\beta\sin\gamma)^{1-2\lambda}\int_{|\beta-\gamma|}^{B}P_n^{(\lambda)}(\cos\alpha)E^{\lambda-1}\sin\alpha\,d\alpha,\end{aligned} \tag{3.4}$$

where

$$B = \begin{cases}\beta+\gamma, & \text{if } \beta+\gamma \leqq \pi,\\ 2\pi-(\beta+\gamma), & \text{if } \beta+\gamma > \pi.\end{cases} \tag{3.5}$$

After this transformation, (3.2) appears in the form

$$\begin{aligned}&\int_{|\beta-\gamma|}^{B}P_n^{(\lambda)}(\cos\alpha)(\sin\alpha\sin\beta\sin\gamma)^{1-2\lambda}E^{\lambda-1}\sin^{2\lambda}\alpha\,d\alpha\\ &\qquad= 2\{\Gamma(\lambda)\}^2\frac{\Gamma(n+1)}{\Gamma(n+2\lambda)}P_n^{(\lambda)}(\cos\beta)P_n^{(\lambda)}(\cos\gamma).\end{aligned} \tag{3.6}$$

This means, according to (2.2), that (1.2) is indeed the "formal" expansion of the function $f(\alpha)$, which is defined by the first expression on the right side of (1.2) if $|\beta-\gamma| < \alpha < B$, and by the second value, namely, zero, if $0 \leqq \alpha < |\beta-\gamma|$ or $B < \alpha \leqq \pi$.

The convergence of this expansion can readily be discussed by means of the asymptotic formula of $P_n^{(\lambda)}(\cos\theta)$ for large values of n [see (4.2)]. The only difficulty is in showing the identity of the sum of this expansion with $f(\alpha)$. This can be concluded for example from the Abel-summability of the ultraspherical expansion of $f(\alpha)$ at an arbitrary point where $f(\alpha)$ is continuous. This follows in the same way as in the case of the ordinary Fourier series and Poisson integral, because the "kernel", namely,

$$\sum_{n=0}^{\infty}(n+\lambda)r^nP_n^{(\lambda)}(\cos A) = \lambda\frac{1-r^2}{(1-2r\cos A+r^2)^{\lambda+1}}, \tag{3.7}$$

is positive and $\to 0$, as $r\to 1$, provided $\cos A \neq 1$ [**3**, p. 71].

4. **Derivation of Sonine's formula** (1.3) **from Dougall's integral formula** (1.1). We base our argument on Mehler's formula [**6**, p. 156],

$$\lim_{n\to\infty} n^{1-2\lambda} P_n^{(\lambda)}(\cos z/n) = \lim_{n\to\infty} n^{-2\nu} P_n^{(\nu+\frac{1}{2})}(\cos z/n)$$
$$(4.1) \qquad = \frac{\pi^{\frac{1}{2}}}{\Gamma(\nu+\frac{1}{2})}(2z)^{-\nu}J_\nu(z) \qquad (\lambda = \nu + \tfrac{1}{2}),$$

and on the following asymptotic formula due to Stieltjes [cf. **2**, p. 302],

$$P_n^{(\lambda)}(\cos\theta) = P_n^{(\nu+\frac{1}{2})}(\cos\theta)$$
$$(4.2) \qquad = \frac{2}{\Gamma(\nu+\frac{1}{2})}(2\sin\theta)^{-\nu-\frac{1}{2}} n^{\nu-\frac{1}{2}}[\cos(n'\theta - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + R],$$
$$n' = n + \nu + \tfrac{1}{2}, \qquad |R| < C(n\sin\theta)^{-1}.$$

The convergence in (4.1) is uniform with respect to z in every bounded region of the complex z-plane, and in particular, in every finite real interval; in (4.2) $0 < \theta < \pi$ and C is a constant independent of n and θ. We write $\lambda = \nu + \frac{1}{2}$ and assume $\lambda > 0$.

Let a, b, c be arbitrary positive numbers, N a positive integer, and[5]

$$(4.3) \qquad l = 2[\tfrac{1}{2}aN], \qquad m = 2[\tfrac{1}{2}bN], \qquad n = 2[\tfrac{1}{2}cN].$$

If a, b, c are the sides of a triangle with a positive area, the same is true for l, m, n, if N is large. Furthermore, if $a + b + c = 2\sigma$, we have $l + m + n = 2s \cong 2\sigma N$, when $N \to \infty$, and therefore according to Stirling's formula,

$$(4.4) \qquad D_1(\lambda; l, m, n) \cong \pi 2^{1-2\lambda}\{\Gamma(\lambda)\}^{-4}[\sigma(\sigma-a)(\sigma-b)(\sigma-c)]^{\lambda-1}N^{4\lambda-4}$$
$$= \pi 2^{-2\nu}\{\Gamma(\nu+\tfrac{1}{2})\}^{-4}\Delta^{2\nu-1}N^{4\nu-2}.$$

If a, b, c are not sides of a triangle, that is, $\max(a, b, c) > s$, the same is true for l, m, n if N is sufficiently large. Thus, we need only to prove, under the conditions $a > 0, b > 0, c > 0, -a + b + c \neq 0, a - b + c \neq 0, a + b - c \neq 0$, that

$$\tfrac{1}{2}2^{3\nu}\pi^{-\frac{1}{2}}\{\Gamma(\nu+\tfrac{1}{2})\}^3(abc)^{-\nu}\lim_{N\to\infty} N^{2-4\nu}D_1(\lambda; l, m, n)$$
$$(4.5) \qquad = \tfrac{1}{2}K\lim_{N\to\infty} N^{2-4\nu}D_1(\lambda; l, m, n) = S(\nu; a, b, c)$$

holds. But since $l + m + n$ is even,

$$(4.6) \qquad \tfrac{1}{2}KN^{2-4\nu}D_1(\lambda; l, m, n) = KN^{2-4\nu}\int_0^{\frac{1}{2}\pi} P_l^{(\lambda)}(\cos\theta)P_m^{(\lambda)}(\cos\theta)P_n^{(\lambda)}(\cos\theta)\sin^{2\lambda}\theta\, d\theta.$$

Let ω be an arbitrarily chosen positive number independent of N; we can break up the interval $[0, \frac{1}{2}\pi]$ of the integration into the two intervals from 0 to ω/N

[5] $[x]$ means the integral part of the number x.

and from ω/N to $\frac{1}{2}\pi$. The integral corresponding to the first of these intervals gives, for $N \to \infty$, according to (4.1),

$$
\begin{aligned}
&KN^{-6\nu} \int_0^\omega P_l^{(\lambda)}(\cos\phi/N)P_m^{(\lambda)}(\cos\phi/N)P_n^{(\lambda)}(\cos\phi/N)\sin^{2\lambda}(\phi/N)N^{2\lambda}\,d\phi \\
&\qquad\qquad \to \int_0^\omega \phi^{1-\nu}J_\nu(a\phi)J_\nu(b\phi)J_\nu(c\phi)\,d\phi.
\end{aligned}
\tag{4.7}
$$

In the integral corresponding to the other interval, we use (4.2) and obtain

$$
\begin{aligned}
&KN^{2-4\nu}P_l^{(\lambda)}(\cos\theta)P_m^{(\lambda)}(\cos\theta)P_n^{(\lambda)}(\cos\theta)\sin^{2\lambda}\theta \\
&\quad = KN^{2-4\nu}2^{-3\nu+\frac{3}{2}}\{\Gamma(\nu+\tfrac{1}{2})\}^{-3}(lmn)^{\nu-\frac{1}{2}}(\sin\theta)^{-\nu-\frac{1}{2}}\{\cos(l'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)+r_1\} \\
&\qquad\qquad \cdot\{\cos(m'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)+r_2\}\{\cos(n'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)+r_3\} \\
&\quad = (2/\pi)^{\frac{3}{2}}(abc)^{-\nu}(lmn)^{\nu-\frac{1}{2}}N^{2-4\nu}(\sin\theta)^{-\nu-\frac{1}{2}}\{\cos(l'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi) \\
&\qquad\qquad \cos(m'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)\cos(n'\theta-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)+r\}.
\end{aligned}
\tag{4.8}
$$

Here for r_1, r_2, r_3, and consequently also for r, bounds of the type given for $|R|$ in (4.2) [with N instead of n] hold. The symbols l', m' have a meaning analogous to that of n' in (4.2).

Now we use the identity

$$
\begin{aligned}
&4\cos(l'\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)\cos(m'\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi)\cos(n'\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi) \\
&\quad = \cos\{(l'+m'+n')\phi-\tfrac{3}{2}\nu\pi-\tfrac{3}{4}\pi\}+\cos\{(m'+n'-l')\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi\} \\
&\quad + \cos\{(n'+l'-m')\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi\}+\cos\{(l'+m'-n')\phi-\tfrac{1}{2}\nu\pi-\tfrac{1}{4}\pi\}.
\end{aligned}
\tag{4.9}
$$

Thus the part $(\omega/N, \frac{1}{2}\pi)$ of the integral corresponding to the principal term of (4.8) can be reduced to expressions of the type

$$
P^{\frac{1}{2}-\nu}\int_{\omega/N}^{\frac{1}{2}\pi}\sin^{-\nu-\frac{1}{2}}\theta\cos(Q\theta+\tau)\,d\theta,
\tag{4.10}
$$

where P and Q are $\sim N$, and τ is a fixed constant. By the second mean value theorem, formula (4.10) becomes

$$
P^{\frac{1}{2}-\nu}(\sin\omega/N)^{-\nu-\frac{1}{2}}O(Q^{-1}) = \omega^{-\nu-\frac{1}{2}}O(1).
\tag{4.11}
$$

This holds uniformly with respect to ω. The contribution of the remainder term is

$$
P^{\frac{1}{2}-\nu}\int_{\omega/N}^{\frac{1}{2}\pi}(\sin\theta)^{-\nu-\frac{1}{2}}(N\sin\theta)^{-1}\,d\theta = N^{-\nu-\frac{1}{2}}(\omega/N)^{-\nu-\frac{1}{2}}O(1) = \omega^{-\nu-\frac{1}{2}}O(1).
\tag{4.12}
$$

Thus, for $\omega \to \infty$, the statement expressed in (4.5) is proved.

5. **Derivation of Sonine's formula** (1.3) **from Dougall's expansion formula** (1.2). Let a, b, c, and h be arbitrary positive numbers, and

$$
\alpha = ah, \qquad \beta = bh, \qquad \gamma = ch.
\tag{5.1}
$$

If a, b, c are the sides of a plane triangle with a positive area Δ, then α, β, γ will form the sides of a spherical triangle, provided h is small. Furthermore, as $h \to 0$, and $\lambda = \nu + \frac{1}{2}$,

$$(5.2)\quad \begin{aligned} D_2(\lambda; \alpha, \beta, \gamma) &\cong 2^{-2\lambda}\pi\{\Gamma(\lambda)\}^{-4}(abc)^{1-2\lambda}h^{-2\lambda-1}\Delta^{2\lambda-2} \\ &= 2^{-2\nu-1}\pi\{\Gamma(\nu + \tfrac{1}{2})\}^{-4}(abc)^{-2\nu}h^{-2\nu-2}\Delta^{2\nu-1}. \end{aligned}$$

If a, b, c are not the sides of a plane triangle, that is, $\max(a, b, c) > \frac{1}{2}(a + b + c)$, then α, β, γ will not form the sides of a spherical triangle, if h is sufficiently small. Thus we need only to prove, under the conditions $a > 0$, $b > 0$, $c > 0$, $-a + b + c \neq 0$, $a - b + c \neq 0$, $a + b - c \neq 0$, that

$$(5.3)\quad \begin{aligned} &2^{3\nu}\pi^{-\frac{3}{2}}\{\Gamma(\nu + \tfrac{1}{2})\}^{3}(abc)^{\nu} \lim_{h\to 0} h^{2\nu+2}D_2(\lambda; \alpha, \beta, \gamma) \\ &\qquad = K_1 \lim_{h \to 0} h^{2\nu+2}D_2(\lambda; \alpha, \beta, \gamma) = S(\nu; a, b, c) \end{aligned}$$

holds. We observe that $K_1 = (abc)^{2\nu}K$.

In order to prove (5.3), we break up the summation, with respect to n, in $D_2(\lambda; \alpha, \beta, \gamma)$ into two ranges; namely, into the ranges $0 \leqq n \leqq [\omega/h]$ and $n \geqq [\omega/h] + 1$, where ω is an arbitrarily chosen positive number, independent of h. Now, according to Mehler's formula (4.1), we have, for $nz = O(1)$,

$$(5.4)\quad n^{-2\nu}P_n^{(\nu+\frac{1}{2})}(\cos z) = \frac{\pi^{\frac{1}{2}}}{\Gamma(\nu + \frac{1}{2})}(2nz)^{-\nu}J_\nu(nz) + \epsilon_n,$$

where $\epsilon_n \to 0$, as $n \to \infty$. Consequently since the right member is bounded,

$$(5.5)\quad \begin{aligned} &n^{-6\nu}P_n^{(\nu+\frac{1}{2})}(\cos \alpha)P_n^{(\nu+\frac{1}{2})}(\cos \beta)P_n^{(\nu+\frac{1}{2})}(\cos \gamma) \\ &\qquad = \pi^{\frac{3}{2}}\{\Gamma(\nu + \tfrac{1}{2})\}^{-3}(2n)^{-3\nu}(\alpha\beta\gamma)^{-\nu}J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) + \eta_n, \end{aligned}$$

where $\eta_n \to 0$, as $n \to \infty$. We observe also that

$$(5.6)\quad (n + \lambda)\left\{\frac{\Gamma(n + 1)}{\Gamma(n + 2\lambda)}\right\}^2 = n^{3-4\lambda}(1 + \eta_n') = n^{1-4\nu}(1 + \eta_n'),$$

where $\eta_n' = O(n^{-1})$ as $n \to \infty$. Thus the summation over the first range contributes[6]

$$(5.7)\quad \begin{aligned} &K_1h^{2\nu+2}\sum_{n=0}^{[\omega/h]}(n + \lambda)\left\{\frac{\Gamma(n + 1)}{\Gamma(n + 2\lambda)}\right\}^2 P_n^{(\lambda)}(\cos \alpha)P_n^{(\lambda)}(\cos \beta)P_n^{(\lambda)}(\cos \gamma) \\ &= K_1h^{2\nu+2}\sum_{n=0}^{[\omega/h]} n^{1-4\nu}(1 + \eta_n') \\ &\qquad\qquad n^{6\nu}\{\pi^{\frac{3}{2}}[\Gamma(\nu + \tfrac{1}{2})]^{-3}(2n)^{-3\nu}(\alpha\beta\gamma)^{-\nu}J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) + \eta_n\} \\ &= h^{2\nu+2}\sum_{n=0}^{[\omega/h]} n^{1-4\nu}(1 + \eta_n')n^{6\nu}(nh)^{-2\nu-1}\{(nh)^{1-\nu}J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) \\ &\qquad\qquad + K_1(nh)^{2\nu+1}\eta_n\} \\ &= h\sum_{n=0}^{[\omega/h]}\{(nh)^{1-\nu}J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) + K_1(nh)^{2\nu+1}\eta_n\}(1 + \eta_n'). \end{aligned}$$

[6] Here and in what follows $n^{-\mu}$, where $\mu \geqq 0$, must be replaced by a constant, for example by 1, for $n = 0$.

The principal part of (5.7)

$$h \sum_{n=0}^{[\omega/h]} (nh)^{1-\nu} J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) \to \int_0^\omega x^{1-\nu} J_\nu(ax)J_\nu(bx)J_\nu(cx)\,dx, \tag{5.8}$$

as $h \to 0$, for $\nu > -\frac{1}{2}$. Since $(nh)^{1-\nu}J_\nu(n\alpha)J_\nu(n\beta)J_\nu(n\gamma) = (nh)^{2\nu+1}O(1)$, the remainder part contributes

$$h \sum_{n=0}^{[\omega/h]} (nh)^{2\nu+1}\eta_n = h^{2\nu+2} \sum_{n=0}^{[\omega/h]} n^{2\nu+1}\eta_n \to 0, \tag{5.9}$$

as $h \to 0$.

In the summation over the second range, we use Stieltjes' formula (4.2) and obtain

$$\begin{aligned} K_1 h^{2\nu+2} n^{1-4\nu}(1 + \eta_n')P_n^{(\nu+\frac{1}{2})}(\cos\alpha)P_n^{(\nu+\frac{1}{2})}(\cos\beta)P_n^{(\nu+\frac{1}{2})}(\cos\gamma) \\ = (2/\pi)^{\frac{3}{2}}(abc)^\nu h^{2\nu+2}(n\sin\alpha\sin\beta\sin\gamma)^{-\nu-\frac{1}{2}}\{\cos(n'\alpha - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + \rho_1\} \\ \cdot\{\cos(n'\beta - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + \rho_2\}\{\cos(n'\gamma - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + \rho_3\} \\ = (2/\pi)^{\frac{3}{2}}(abc)^\nu h^{2\nu+2}(n\sin\alpha\sin\beta\sin\gamma)^{-\nu-\frac{1}{2}}\{\cos(n'\alpha - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) \\ \cos(n'\beta - \tfrac{1}{2}\pi\nu - \tfrac{1}{4}\pi)\cos(n'\gamma - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + \rho\}. \end{aligned} \tag{5.10}$$

Here for ρ_1, ρ_2, ρ_3, and consequently also for ρ, bounds of the type given for $|R|$ in (4.2) hold. We have $\rho = (nh)^{-1}O(1)$. We also used $\eta_n' = O(n^{-1})$.

By means of the identity (4.9), the principal term of (5.10) can be reduced to an expression of the type

$$p^{\frac{1}{2}-\nu} \sum_{n=[\omega/h]+1}^{\infty} n^{-\nu-\frac{1}{2}} \cos(nq + \tau_1), \tag{5.11}$$

where p and q are $\sim h$, and τ_1 is a fixed constant. By use of Abel's transformation we find for (5.11) the bound

$$p^{\frac{1}{2}-\nu}(h^{-1}\omega)^{-\nu-\frac{1}{2}}q^{-1} = \omega^{-\nu-\frac{1}{2}}O(1). \tag{5.12}$$

The contribution of the remainder term is

$$p^{\frac{1}{2}-\nu} \sum_{n=[\omega/h]+1}^{\infty} n^{-\nu-\frac{1}{2}}(nh)^{-1} = O(1)h^{-\nu-\frac{1}{2}} \sum_{n=[\omega/h]+1}^{\infty} n^{-\nu-\frac{3}{2}} = \omega^{-\nu-\frac{1}{2}}O(1). \tag{5.13}$$

Recapitulating, we obtain, for $h \to 0$,

$$K_1 h^{2\nu+2} D_2(\lambda; \alpha, \beta, \gamma) = \int_0^\omega x^{1-\nu} J_\nu(ax)J_\nu(bx)J_\nu(cx)\,dx + o(1) + \omega^{-\nu-\frac{1}{2}}O(1). \tag{5.14}$$

Thus, for $\omega \to \infty$, the statement expressed in (5.3) is proved.

References

1. J. Dougall, *A theorem of Sonine in Bessel functions, with two extensions to spherical harmonics*, Proc. Edin. Math. Soc., vol. 37(1919), pp. 33–47.
2. E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Cambridge, 1931.
3. E. Kogbetliantz, *Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typiques*, Mémorial des Sciences Mathématiques, Paris, Fascicule LI (1931).
4. G. Szegö, *Asymptotische Entwicklungen der Jacobischen Polynome*, Schriften der Königsberger Gelehrten Gesellschaft, Naturw. Klasse (1933), pp. 35–111.
5. I. Todhunter, *Note on Legendre's coefficients*, Proc. Roy. Soc., vol. 27(1878), pp. 381–383.
6. G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge, 1922.
7. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th Edition, Cambridge, 1935.

Washington University.

TAYLOR'S SERIES OF FUNCTIONS OF SMOOTH GROWTH IN THE UNIT CIRCLE

By W. T. Martin and N. Wiener

1. **Introduction.** We have recently proved a series of results connecting the growth of an entire function with the growth of certain expressions depending on the coefficients.[1] There is a closely related series of results which concern, not the growth of an entire function, but the growth of a function as it approaches the circle of convergence. As will be seen in the formulation of our results, there is a precise methodological similarity between theorems of the two types, and, indeed, they differ merely by the fact that a parameter which is negative in one case is positive in the other. It is nevertheless worth while to give an explicit formulation of the theorems because they are of less familiar character than the other set of theorems and involve results which have only recently been obtained by Vijayaraghavan and Wiener, Avakumović and Karamata, and Pitt (see footnotes 4, 6) in contrast with the theorems of the entire type which are best exemplified by the classical theory of the Borel summation. The circle of convergence theorems are more directly applicable to a series of interesting problems in the analytical theory of numbers, and, in particular, they allow us to carry Tauberian methods in the problem of partitions much further than Hardy and Ramanujan at one time believed possible.[2] There is a considerable prospect of their further utility in unexplored portions of this field. It is true that the revolutionary work of Rademacher[3] overshadows all less perfect methods, whether Tauberian or not. We wish to call attention to the fact that it is possible by an artifice to eliminate the condition of positivity which is necessary for Tauberian theorems of the type considered, or, more accurately, to guarantee its satisfaction by a consideration not of a singularity of a power series at an individual point on the circle of convergence, but rather the behavior of the integral of the square of its modulus as we approach the circle of convergence.

There will be many places in which the detail of proof is so similar to that of our previous paper that anything like completeness is scarcely necessary. Where this is the case we shall consider ourselves at liberty to present our argument in a highly schematic form and to refer to our previous paper for technical details.

Received January 12, 1938.

[1] N. Wiener and W. T. Martin, *Taylor's series of entire functions of smooth growth*, this Journal, vol. 3(1937), pp. 213-223.

[2] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. of the London Math. Soc., vol. 17(1918), pp. 75-115. See especially pp. 89, 90.

[3] H. Rademacher, *On the partition function* $p(n)$, Proc. of the London Math. Soc., vol. 43 (1937), pp. 241-254.

Let $a_n \geqq 0$ ($n = 1, 2, \cdots$), and let

$$\sum_1^\infty a_n x^n \sim sH(x) \qquad (x \to 1^-). \tag{1.1}$$

We may rewrite (1.1) in the form

$$\sum_1^\infty a_n e^{F(\xi)-n\xi} \to s \qquad (\xi \to 0^+), \tag{1.2}$$

where

$$F(\xi) = -\log H(e^{-\xi}). \tag{1.3}$$

We shall derive the following Tauberian theorem.

THEOREM 1. *Let $a_n \geqq 0$ ($n = 1, 2, \cdots$), and let* (1.2) *hold, where*

(1.4) *F is four times continuously differentiable for $0 < \xi \leqq \xi_0$ for some positive ξ_0;*

$$F''(\xi) \leqq \text{const.} < 0 \quad \textit{for } 0 < \xi \leqq \xi_0; \tag{1.5}$$

$$\int_\xi^{\xi_0} [-F''(t)]^{\frac{1}{2}}\, dt \to \infty \qquad (\xi \to 0^+); \tag{1.6}$$

$$F'''(\xi) = o([-F''(\xi)]^{\frac{3}{2}}) \qquad (\xi \to 0^+); \tag{1.7}$$

$$F^{\text{iv}}(\xi) = o([F''(\xi)]^2) \qquad (\xi \to 0^+). \tag{1.8}$$

Then for every positive λ,

$$\lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq \psi(n) < x+\lambda} a_n e^{F(G(n))-nG(n)} = s, \tag{1.9}$$

where $\psi(x)$ is defined by

$$\int_a^x [-G'(x)]^{\frac{1}{2}}\, dx = \psi(x) \tag{1.10}$$

and $G(x)$ (for $x > a = F'(\xi_0)$) is the inverse function to $F'(\xi)$; that is, $G(F'(\xi)) = \xi$ for $0 < \xi \leqq \xi_0$.

Theorem 1 and Theorem 4, which we will deduce from Theorem 1, have important applications in the analytical theory of numbers. In a later paper we hope to present some results based upon them.

A special case of Theorem 1, due to Vijayaraghavan and Wiener (in an unpublished paper), is the following

THEOREM A. *Let $a_n \geqq 0$ ($n = 1, 2, \cdots$), and let*

$$\sum_1^\infty a_n x^n \sim s \exp[(1 - x)^{-1}] \qquad (x \to 1^-). \tag{1.11}$$

Then for every positive λ

$$\lim_{N\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{N \leqq n < N+\lambda N^{\frac{1}{2}}} a_n \exp\left[-2n^{\frac{1}{2}}\right] = s\,\frac{e^{\frac{1}{2}}}{2^{\frac{1}{2}}}. \tag{1.12}$$

To derive Theorem A from Theorem 1 we write $x = e^{-\xi}$. Then (1.11) becomes

$$\sum a_n e^{-n\xi} \sim s \exp\left\{\frac{1}{1-e^{-\xi}}\right\} = s e^{1/\xi} \exp\left\{\frac{1}{1-e^{-\xi}} - \frac{1}{\xi}\right\}. \tag{1.13}$$

Noting that

$$\exp\left\{\frac{1}{1-e^{-\xi}} - \frac{1}{\xi}\right\} \to e^{\frac{1}{2}} \quad \text{as } \xi \to 0^+,$$

we see that (1.13) becomes

$$\sum a_n \exp\left\{-\frac{1}{\xi} - n\xi\right\} \to s e^{\frac{1}{2}} \qquad (\xi \to 0^+). \tag{1.14}$$

Thus for this case $F(\xi) = -\xi^{-1}$, and it is easily verified that this function satisfies conditions (1.4), $\cdots$, (1.8). Moreover

$$G(x) = x^{-\frac{1}{2}}, \qquad \psi(x) = 2^{\frac{1}{2}} x^{\frac{1}{2}}$$

and (1.9) becomes

$$\lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq 2^{\frac{1}{2}} n^{\frac{1}{2}} < x+\lambda} a_n \exp\left[-2n^{\frac{1}{2}}\right] = s e^{\frac{1}{2}} \tag{1.15}$$

from which (1.12) easily follows.

Avakumović and Karamata have proved the following[4]

THEOREM B. *Let* $A(y)$ *be* $\geqq 0$ *and non-decreasing for* $y \geqq 0$ *and let*

$$J(x) = \int_0^\infty e^{-y/x} A(y)\, dy$$

converge for $x > 0$. *Let*[5]

$$J(x) \asymp x^k e^x \qquad (x \to \infty) \tag{1.16}$$

for some real k. *Then*

$$\exp\left[2x^{\frac{1}{2}} - (2+\epsilon)x^{\frac{1}{4}}(\log x)^{\frac{1}{2}}\right] < A(x) < M x^{\frac{1}{2}(k-1)} \exp\left[2x^{\frac{1}{2}}\right]. \tag{1.17}$$

H. R. Pitt has proved a theorem which generalizes Theorem B, at the same time furnishing more precise results.[6] His theorem in this particular case is

[4] Avakumović and J. Karamata, *Über einige Taubersche Sätze, deren Asymptotik von Exponentialcharacter ist.* I, Mathematische Zeitschrift, vol. 41(1936), pp. 345–356.

[5] Hardy writes $f(x) \asymp g(x)$ $(x \to \infty,\ g(x) > 0)$ if there are two constants $0 < \delta < \Delta < \infty$ such that $\delta < f(x)/g(x) < \Delta$ for $x > x_0$.

[6] Pitt discovered and proved his theorem independently of our results. In conversation with him we discussed the results and later he showed how the case $k = 0$ of his theorem is deducible from ours. In §3 we shall give the proof of this portion of Pitt's theorem as submitted to us by him.

THEOREM C. *Suppose that $A(y)$ is non-decreasing in $(0, \infty)$ and that*

$$J(x) = \int_0^\infty e^{-y/x}\, dA(y)$$

converges absolutely for $x > 0$. Let k be any real number,

$$\alpha > 0, \qquad B = \alpha^{-\alpha/(1+\alpha)}(1 + \alpha),$$

$$h(x) = e^{-Bx^2} x^{1-2k/\alpha} A(x^{2(1+\alpha)/\alpha}).$$

Let

$$J(x) \sim x^k \exp [x^\alpha] \qquad (x \to \infty).$$

Then

$$e^{-\epsilon(x)} \leqq h(x) \leqq \epsilon(x), \tag{1.18}$$

where

$$\lim_{x\to\infty} \frac{\epsilon(x)}{x} = 0.$$

As a special case of Theorem C Pitt has obtained the following extension of Avakumović's and Karamata's result.

THEOREM D. *If $A(y)$ is non-decreasing in $(0, \infty)$ and*

$$J(x) \sim x^k e^x \qquad (x \to \infty),$$

then

$$\exp [2x^{\frac{1}{2}} - o(x^{\frac{1}{2}})] \leqq A(x) \leqq o(x^{\frac{1}{2}k} \exp [2x^{\frac{1}{2}}]). \tag{1.19}$$

As a converse to Theorem 1 we shall prove

THEOREM 2. *Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, and let* (1.9) *hold for every positive value of λ, where F fulfills the conditions* (1.4), $\cdots$, (1.8) *and ψ and G are defined as in Theorem* 1. *Then* (1.2) *holds.*

In the next theorem we place some additional restrictions on F and we obtain certain inequalities related to ψ.

THEOREM 3. (i) *Let F fulfill conditions* (1.4), (1.5), (1.6) *and the additional condition:*

(1.20) *for some p in $\frac{1}{2} < p < 1$ and for all sufficiently small positive ϵ,*

$$-F''(\xi) = O([F'(\xi)]^{2p+\epsilon}) \qquad (\xi \to 0^+).$$

Then for every positive ϵ

$$n^{-p-\epsilon} < \psi'(n) \tag{1.21}$$

and

$$n^{1-p-\epsilon} < \psi(n) \tag{1.22}$$

for sufficiently large values of n.

(ii) *Let F fulfill conditions* (1.4), (1.5), (1.6) *and the additional condition*:

(1.23) *for some p in $\frac{1}{2} < p < 1$ and for all sufficiently small positive ϵ,*

$$[F'(\xi)]^{2p-\epsilon} < -F''(\xi)$$

for sufficiently small (positive) values of ξ.
Then for every positive ϵ,

$$\psi(n) < n^{1-p+\epsilon} \tag{1.24}$$

for sufficiently large values of n.

As immediate consequences of these theorems we shall obtain the following theorems on smoothly growing functions analytic in the unit circle.

THEOREM 4. *Let*

$$f(z) = \sum_0^\infty b_n z^n \tag{1.25}$$

be analytic for $|z| < 1$, and let

$$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\vartheta})|^2 \, d\vartheta \sim se^{-F(-2\log r)} \qquad (r \to 1^-), \tag{1.26}$$

where F fulfills conditions (1.4), $\cdots$, (1.8). *Then for every positive λ,*

$$\lim_{x \to \infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leq \psi(n) < x+\lambda} |b_n|^2 e^{F(G(n)) - nG(n)} = s, \tag{1.27}$$

where ψ and G are defined as in Theorem 1.

The converse to Theorem 4 is

THEOREM 5. *Let* (1.25) *be analytic for $|z| < 1$, and let* (1.27) *hold for every positive value of λ, where F fulfills conditions* (1.4), $\cdots$, (1.8) *and ψ and G are defined as in Theorem* 1. *Then* (1.26) *holds.*

If we place the additional restrictions (1.20) and (1.23) on F, we have the following gap theorem.

THEOREM 6. *Let* (1.25) *be analytic for $|z| < 1$, and let* (1.26) *hold, where $s \neq 0$ and F satisfies conditions* (1.4), $\cdots$, (1.8), (1.20), (1.23). *Then* (1.27) *holds where ψ and G are defined as in Theorem* 1. *Furthermore, for every positive ϵ we shall always have* (1.21) *and*

$$n^{1-p-\epsilon} < \psi(n) < n^{1-p+\epsilon} \tag{1.28}$$

for sufficiently large values of n. Thus the function (1.25) *can have only a finite number of gaps of magnitude $(\nu, \nu + \nu^{p+\epsilon})$, that is, for any positive ϵ, the equations*

$$a_n = a_{n+1} = \cdots = a_{n+[n^{p+\epsilon}]} = 0$$

can hold for at most a finite number of values of n.

A slightly more general form of Theorem 1, which we shall first prove, is

THEOREM 7. *Let* $\{\lambda_n\}$ *be any sequence such that*

$$(1.29)\qquad 0 \leqq \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots, \qquad \lim_{n\to\infty} \lambda_n = \infty;$$

let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, *and let*

$$(1.30)\qquad \sum_1^\infty a_n e^{F(\xi)-\lambda_n\xi} \to s \qquad (\xi \to 0^+),$$

where F *satisfies conditions* (1.4), $\cdots$, (1.8). *Then for every positive* λ,

$$(1.31)\qquad \lim_{x\to\infty} \frac{(2\pi)^{\frac{1}{2}}}{\lambda} \sum_{x \leqq \psi(\lambda_n) < x+\lambda} a_n e^{F(G(\lambda_n))-\lambda_n G(\lambda_n)} = s,$$

where ψ *and* G *are defined as in Theorem* 1. *Conversely, if* $\{\lambda_n\}$ *is any sequence satisfying* (1.29), *if* $a_n \geqq 0$ $(n = 1, 2, \cdots)$, *and if* (1.31) *holds for every positive value of* λ, *then* (1.30) *holds.*

2. For the proof of Theorem 7 let us consider the exponent in (1.30) with a general argument w instead of λ_n: $F(\xi) - w\xi$. Let us place

$$(2.1)\qquad \xi = \varphi(u + \psi(w)).$$

Developing the exponent $F(\xi) - w\xi$ in a Taylor's series with remainder and determining φ and ψ in such a manner that the coefficient of u vanishes and that the coefficient of $\frac{1}{2}u^2$ is negative unity, we find

$$(2.2)\qquad \psi(w) = \int_a^w [-G'(x)]^{\frac{1}{2}}\,dx, \qquad \varphi(w) = G(\psi^{-1}(w)),$$

where $G(x)$ is the function inverse to $F'(\xi)$: $G(F'(\xi)) = \xi$. The details here are entirely analogous to those of our earlier paper (see reference in footnote 1). With these functions φ, ψ, using the conditions on F, we find that the remainder in the Taylor's series approaches zero; indeed, it approaches zero in such a manner that (1.31) implies that

$$(2.3)\qquad \sum_1^\infty a_n e^{F(G(\lambda_n))-\lambda_n G(\lambda_n)} \to s \qquad (\xi \to 0^+).$$

It is an easy step from this to apply Wiener's General Tauberian Theorem[7] to (2.3). Here the kernel $K(u)$ is $e^{-\frac{1}{2}u^2}$ whose Fourier transform $e^{-\frac{1}{2}x^2}$ is non-vanishing, and u occurs as the difference of two variables: $u = \varphi^{-1}(\xi) - \psi(w)$. The relation (1.31) follows at once.

The second part of Theorem 7 follows at once from the converse portion of Wiener's Tauberian Theorem (see reference in footnote 7, p. 26).

Theorems 1 and 2 are special cases of Theorem 7 for $\lambda_n = n$. Theorems 4 and 5 follow at once from these theorems.

[7] N. Wiener, *Tauberian theorems*, Ann. of Math., (2), vol. 33(1932), pp. 1–100.

Using the definition (2.2) of $\psi(w)$ and the hypotheses on F in Theorem 3, one easily obtains the conclusions of Theorem 3. Theorem 6 follows immediately from Theorems 3 and 4.

3. In this section we shall show how the portion of Pitt's theorem (Theorem C of this paper) for which $k = 0$ is deducible from Theorem 1. The method used in showing this is essentially due to Pitt. Using his method, we first prove the following result from which we will deduce the desired result.

THEOREM E. *Let $A(y)$ be non-decreasing for $0 < y < \infty$, and let $F(\xi)$ satisfy conditions* (1.4), $\cdots$, (1.8). *If*

$$\int_0^\infty e^{-y\xi}\,dA(y) \sim se^{-F(\xi)} \qquad (\xi \to 0^+), \tag{3.1}$$

then for every positive λ there is an x_λ such that

$$(2\pi)^{-\frac{1}{2}}\tfrac{1}{2}s\lambda e^{-R(g(x-\lambda))} \leqq A(g(x)) \leqq h_\lambda(x-\lambda)e^{-R(g(x))} + H_\lambda \sum_{1 \leqq n < x/\lambda} e^{-R(g(x-n\lambda))} \tag{3.2}$$

for $x \geqq x_\lambda$, where

$$R(x) = F(G(x)) - xG(x), \tag{3.3}$$

$$g(x) = \psi^{-1}(x) \qquad (\psi(g(x)) \equiv x), \tag{3.4}$$

$$h_\lambda(x) = \int_{g(x)}^{g(x+\lambda)} e^{R(y)}\,dA(y), \qquad H_\lambda = \operatorname*{l.u.b.}_{0 \leqq x < \infty} h_\lambda(x), \tag{3.5}$$

and G and ψ are defined as in Theorem 1.

As a consequence of Theorem 1 we have for every positive λ

$$\lim_{x\to\infty} h_\lambda(x) = s(2\pi)^{-\frac{1}{2}}\lambda. \tag{3.6}$$

Furthermore, since

$$R'(y) = F'(G(y))G'(y) - yG'(y) - G(y) = -G(y) < 0,$$

it follows that $R(y)$ is decreasing. Since $R(y)$ is decreasing and $A(y)$ is non-decreasing, we have from (3.5)

$$e^{R(g(x+\lambda))}\int_{g(x)}^{g(x+\lambda)} dA(y) \leqq h_\lambda(x) \leqq e^{R(g(x))}\int_{g(x)}^{g(x+\lambda)} dA(y). \tag{3.7}$$

As a matter of convenience let us take $A(0) = 0$. Then from (3.7)

$$\begin{aligned} e^{R(g(x+\lambda))}[A(g(x+\lambda)) - A(g(x))] &\leqq h_\lambda(x) \\ &\leqq e^{R(g(x))}\int_0^{g(x+\lambda)} dA(y) = e^{R(g(x))}A(g(x+\lambda)). \end{aligned} \tag{3.8}$$

Using (3.6) and the second inequality in (3.8), we have

$$\begin{aligned} A(g(x)) &\geqq h_\lambda(x-\lambda)e^{-R(g(x-\lambda))} \\ &\geqq \tfrac{1}{2}s(2\pi)^{-\frac{1}{2}}\lambda e^{-R(g(x-\lambda))} \qquad (x \geqq x_\lambda). \end{aligned} \tag{3.9}$$

The first inequality in (3.8), together with (3.6), gives

$$\begin{aligned} A(g(x)) &\leqq \sum_{0\leqq n<x/\lambda} [A(g(x-n\lambda)) - A(g(x-\overline{n+1}\lambda)) \\ &\leqq \sum_{0\leqq n<x/\lambda} h_\lambda(x-\overline{n+1}\lambda)e^{-R(g(x-n\lambda))} \\ &\leqq h_\lambda(x-\lambda)e^{-R(g(x))} + \sum_{1\leqq n<x/\lambda} h_\lambda(x-\overline{n+1}\lambda)e^{-R(g(x-n\lambda))} \\ &\leqq h_\lambda(x-\lambda)e^{-R(g(x))} + H_\lambda \sum_{1\leqq n<x/\lambda} e^{-R(g(x-n\lambda))}. \end{aligned} \tag{3.10}$$

This concludes the proof of Theorem E.

We now interpret this result for the case

$$F(\xi) = -\xi^{-\alpha}, \qquad \alpha > 0; \qquad s \neq 0. \tag{3.11}$$

Then

$$\begin{aligned} F'(\xi) &= \alpha\xi^{-\alpha-1}, \qquad G(x) = \frac{\alpha^{1/(\alpha+1)}}{x^{1/(\alpha+1)}}, \\ \psi(x) &= 2(\alpha+1)^{\frac{1}{2}}\alpha^{-\frac{1}{2}}\alpha^{-\frac{1}{2}\alpha/(\alpha+1)}x^{\frac{1}{2}\alpha/(\alpha+1)} \\ &= 2\alpha^{-\frac{1}{2}}B^{\frac{1}{2}}x^{\frac{1}{2}\alpha/(\alpha+1)} \qquad (B = (\alpha+1)\alpha^{-\alpha/(\alpha+1)}), \\ R(x) &\equiv F(G(x)) - xG(x) = -Bx^{\alpha/(\alpha+1)}, \\ g(x) &= \left(\frac{\alpha x^2}{4B}\right)^{(\alpha+1)/\alpha}. \end{aligned} \tag{3.12}$$

The relation (3.2) becomes

$$\begin{aligned} (2\pi)^{-\frac{1}{2}}\tfrac{1}{2}s\lambda \exp\left[B\frac{\alpha}{4B}(x-\lambda)^2\right] &\leqq A\left(\left(\frac{\alpha x^2}{4B}\right)^{(\alpha+1)/\alpha}\right) \\ \leqq h_\lambda(x-\lambda)\exp\left[B\frac{\alpha}{4B}x^2\right] &+ H_\lambda \sum_{1\leqq n<x/\lambda} \exp\left[B\frac{\alpha}{4B}(x^2 - 2n\lambda x + n^2\lambda^2)\right]. \end{aligned} \tag{3.13}$$

The first inequality gives

$$A(x^{2(\alpha+1)/\alpha})e^{-Bx^2} \geqq (2\pi)^{-\frac{1}{2}}\tfrac{1}{2}s\lambda \exp[-\alpha^{\frac{1}{2}}B^{\frac{1}{2}}\lambda x], \tag{3.14}$$

$$\liminf_{x\to\infty} x^{-1}[\log A(x^{2(\alpha+1)/\alpha}) - Bx^2] \geqq -B^{\frac{1}{2}}\lambda. \tag{3.15}$$

Since this is true for every positive λ,

$$\liminf_{x\to\infty} x^{-1}[\log A(x^{2(\alpha+1)/\alpha}) - Bx^2] \geqq 0. \tag{3.16}$$

Hence

$$A(x^{2(\alpha+1)/\alpha})e^{-Bx^2} \geqq e^{-o(x)}. \tag{3.17}$$

The second inequality in (3.13) gives

$$\begin{aligned} \exp\left[-B\frac{\alpha}{4B}x^2\right]A\left(\left(\frac{\alpha x^2}{4B}\right)^{(\alpha+1)/\alpha}\right) &\leqq h_\lambda(x-\lambda) + H_\lambda \sum_{1\leqq n<x/\lambda} e^{-\frac{1}{4}\alpha n\lambda(2x-n\lambda)} \\ &\leqq h_\lambda(x-\lambda) + H_\lambda \sum_{1\leqq n<x/\lambda} e^{-\frac{1}{4}\alpha n\lambda x} \\ &\leqq h_\lambda(x-\lambda) + H_\lambda \frac{e^{-\frac{1}{4}\alpha\lambda x} - e^{-\frac{1}{4}\alpha\lambda x[x/\lambda]}}{1 - e^{-\frac{1}{4}\alpha\lambda x}}. \end{aligned} \tag{3.18}$$

Hence

$$\limsup_{x\to\infty} \exp\left[-B\frac{\alpha}{4B}x^2\right]A\left(\left(\frac{\alpha x^2}{4B}\right)^{(\alpha+1)/\alpha}\right) \leqq s(2\pi)^{-\frac{1}{2}}\lambda. \tag{3.19}$$

Since λ is arbitrary,

$$\limsup_{x\to\infty} e^{-Bx^2}A(x^{2(\alpha+1)/\alpha}) \leqq 0. \tag{3.20}$$

Hence

$$e^{-Bx^2}A(x^{2(\alpha+1)/\alpha}) = o(1). \tag{3.21}$$

The relations (3.17) and (3.21) yield the conclusion of Pitt's theorem for $k = 0$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

DEGREE OF APPROXIMATION BY POLYNOMIALS IN z AND $1/z$

BY W. E. SEWELL

1. **Introduction.** A polynomial of degree n in z and $1/z$ is a function of the form

$$(1.10) \qquad r_n(z) = a_{-n}z^{-n} + a_{-n+1}z^{-n+1} + \cdots + a_{-1}z^{-1} + a_0 + \cdots + a_nz^n;$$

we do not assume a_{-n} or a_n different from zero. Riesz[1] has shown that $|r_n(z)| \leqq M$ on $C: |z| = 1$ implies $|r_n'(z)| \leqq Mn$, $|z| = 1$. In this paper we extend this result to various types of Jordan curves (see §2) for a generalized derivative (see §3) of an arbitrary positive order α. In fact, we prove that *if C is a Jordan curve containing the origin in its interior, then $|r_n(z)| \leqq M$, for z on C, implies*[2] $|r_n^\alpha(z)| \leqq MK(\alpha, C)n^{\alpha u}$, $\alpha > 0$, $1 \leqq u \leqq 2$, *where K is a constant depending only on α and C, and u is a constant depending only on C.*

Also let $f(z)$ be defined on C and suppose $|f(z) - r_n(z)| \leqq \epsilon_n$, z on C ($n = 1, 2, \cdots$). If $f(z)$ is continuous on C, there exists[3] for each n a polynomial $r_n(z)$ such that ϵ_n approaches zero as n becomes infinite. Here we study the relation between ϵ_n and the continuity properties of $f(z)$ on C. For an analytic Jordan curve C (see §4) the method consists in mapping the interior of C conformally on $|w| < 1$ and applying results on trigonometric approximation due to de la Vallée Poussin[4] and Jackson.[5] We prove, for example, that, *for C an analytic Jordan curve, the existence of $r_n(z)$ ($n = 1, 2, \cdots$) such that $|f(z) - r_n(z)| \leqq Mn^{-\alpha}$, z on C, $0 < \alpha < 1$, α and M independent of n and z, implies that $f(z)$ satisfies a Lipschitz condition*[6] *of order α on C, and, conversely, $f(z)$ satisfying a Lipschitz condition of order α on C implies the existence of $r_n(z)$ such that $|f(z) - r_n(z)| \leqq Mn^{-\alpha}$, z on C.* For $f(z)$ the boundary function of a function

Received January 26, 1938.

[1] M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23(1914), pp. 354–368.

[2] $f^\alpha(z)$ denotes the generalized derivative of order α of $f(z)$.

[3] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 20, 1935; see p. 38.

[4] Ch.-J. de la Vallée Poussin, *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris, 1919.

[5] Dunham Jackson, *The Theory of Approximation*, American Mathematical Society Colloquium Publications, vol. 11, 1930.

[6] The function $f(z)$ satisfies a Lipschitz condition of order α on C if for z_1 and z_2 arbitrary points on C we have $|f(z_1) - f(z_2)| \leqq L|z_1 - z_2|^\alpha$, where L is a constant independent of z_1 and z_2.

analytic in C the problem has been studied by the author,[7] Curtiss,[8] and Walsh and Sewell.[9]

For more general curves (see §5) we use the methods of the author (SIII) and the results herein on $r_n^\alpha(z)$. We show, for example, that *if $f(z)$ is defined on an arbitrary Jordan curve C containing the origin in its interior, and if $r_n(z)$ $(n = 1, 2, \cdots)$ exists such that $|f(z) - r_n(z)| \leqq Mn^{-\alpha}$, z on C, $\alpha > 2$, α and M independent of n and z, then $f(z)$ has a bounded first derivative on C.*

2. Preliminary definitions and theorems.

Let $w = \phi(z)$, whose inverse is $z = \Phi(w)$, map the interior of C conformally on $|w| < 1$, so that $z = 0$ goes into $w = 0$. We denote by I_R the image under this mapping of the circle $|w| = R < 1$; I_R is an interior level curve. Let $w = \psi(z)$, whose inverse is $z = \Psi(w)$, map the exterior of C conformally on $|w| > 1$, so that $z = \infty$ goes into $w = \infty$. We denote by E_ρ the image of the circle $|w| = \rho > 1$ under this mapping; E_ρ is an exterior level curve. Let P be a point of C, let $d(P, I_R)$ be the greatest lower bound of the distances from P to the points of I_R, and let $d(C, I_R)$ be the greatest lower bound of $d(P, I_R)$ as P traverses C. We take the following definitions and theorems from SIII.

DEFINITION 2.1. *If*

$$0 < N_1 \leqq \left| \frac{\Phi(w_1) - \Phi(w_2)}{w_1 - w_2} \right| \leqq N_2 < \infty, \tag{2.10}$$

uniformly for $|w_1| \leqq 1$, $|w_2| \leqq 1$, we shall say that "C is a curve of type S".

DEFINITION 2.2. *Let C be a Jordan curve composed of a finite number of Jordan arcs meeting in corners $z_1, z_2, \cdots, z_p$, of exterior openings $\mu_1\pi, \mu_2\pi, \cdots, \mu_p\pi$, $2 > \mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_p > 0$, and let the difference-quotient of the mapping functions $w = \psi(z)$, $w = \phi(z)$ be bounded in modulus on each smooth subarc. Let u be the larger of μ_1 and $2 - \mu_p$. Then we shall say that "C is a curve of type u".*

DEFINITION 2.3. *Let C be a rectifiable Jordan curve. Let z_1 and z_2 be arbitrary points, $z_1 \neq z_2$, of C. Let z_3 be a point of C distinct from z_1 and z_2, and choose the direction on C so that z_2 separates z_1 and z_3. If constants N_1 and N_2, independent of z_1, z_2 and z_3, exist such that*

$$\int_{z_1}^{z_2} |z_3 - x|^{-\beta} |dx| \leqq N_1 |z_2 - z_3|^{-\beta+1} + N_2, \qquad \beta > 0, \tag{2.11}$$

[7] W. E. Sewell, (I) *Degree of approximation by polynomials to continuous functions*, Bulletin of the American Mathematical Society, vol. 41(1935), pp. 111–117; (II) *On the modulus of the derivative of a polynomial*, ibid., vol. 42(1936), pp. 699–701; (III) *Generalized derivatives and approximation by polynomials*, Transactions of the American Mathematical Society, vol. 41(1937), pp. 84–123. These papers will be referred to as SI, SII, and SIII, respectively.

[8] John Curtiss, *A note on the degree of polynomial approximation*, Bulletin of the American Mathematical Society, vol. 42(1936), pp. 873–878.

[9] J. L. Walsh and W. E. Sewell, *Note on the relation between continuity and degree of polynomial approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 43(1937), pp. 557–563.

where the path of integration is along C from z_1 to z_2, for z_2 arbitrarily near z_3, we shall say that "C is a curve of type W'".

THEOREM 2.4. *Let C be a curve of type S. Then we have*

$$d(C, E_\rho) \geqq M_1(\rho - 1), \qquad d(C, I_R) \geqq M_2(1 - R), \tag{2.12}$$

where M_1 and M_2 are constants depending only on C.

THEOREM 2.5. *Let C be a curve of type u. Then we have*

$$d(C, E_\rho) \geqq M_1(\rho - 1)^u, \qquad d(C, I_R) \geqq M_2(1 - R)^u, \qquad 1 < u < 2, \tag{2.13}$$

where M_1 and M_2 are constants depending only on C.

THEOREM 2.6. *Let C be an arbitrary rectifiable Jordan curve. Then we have*

$$d(C, E_\rho) \geqq M_1(\rho - 1)^2, \qquad d(C, I_R) \geqq M_2(1 - R)^2, \tag{2.14}$$

where M_1 and M_2 are constants depending only on C.

3. **Generalized derivatives of** $r_n(z)$. Let $f(z)$ be defined and integrable on C. Then the Riemann-Liouville generalized derivative of $f(z)$ (see SIII), if it exists, is defined at a point z of C as follows:[10]

$$\begin{aligned} &D_z^0 f(z) \equiv f(z); \\ &D_z^\alpha f(z) = \frac{1}{\Gamma(-\alpha)} \int_k^z (z - x)^{-\alpha-1} f(x)\, dx, && \alpha < 0; \\ &D_z^\alpha f(z) = \frac{d^p}{dz^p} D_z^{\alpha-p} f(z), && 0 \leqq p - 1 \leqq \alpha < p, \end{aligned} \tag{3.10}$$

where p is a positive integer. Here the point k is an arbitrary but fixed point on C and the path of integration is along C in a fixed direction. For analytic functions and $\alpha > 0$ a more convenient representation of the generalized derivative is furnished by the following (see SIII):

THEOREM 3.1. *Let $f(z)$ be analytic in the interior of a rectifiable Jordan curve γ which passes through k and contains in its interior the point z and an arc λ of type W' joining k to z, and let $f(z)$ be continuous in the closed limited region bounded by γ. Then we have*

$$D_z^\alpha f(z) = \frac{\Gamma(\alpha + 1)}{2\pi i} \int_\gamma \frac{f(t)\, dt}{(t - z)^{\alpha+1}}, \qquad \alpha > 0, \tag{3.11}$$

where the branch of $(t - z)^{-\alpha-1}$ is determined on λ.

Let k be a fixed point on C and denote by C^m the set of points z on C with $|z - k| > m > 0$; our inequalities hold in general for z on C^m. Let z be an arbitrary point on C^m, and with z as center draw a circle δ of radius so small that δ does not contain the point k or the origin in its closed interior (the precise radius to be determined later), and let A be the first point of intersection of

[10] A particular branch of $(z - x)^{-\alpha-1}$ is understood here (see SIII, pp. 94–96).

δ and C in traversing C from k to z. Then if γ is the path from k to A, A around δ and back to A, and from A along C to k, we have

$$r_n^\alpha(z) = \frac{\Gamma(\alpha + 1)}{2\pi i} \int_\gamma \frac{r_n(t)\, dt}{(t - z)^{\alpha+1}}, \qquad \alpha > 0. \tag{3.12}$$

Thus,

$$| r_n^\alpha(z) | \leqq \frac{\Gamma(\alpha + 1)}{2\pi} \left\{ 2 \left| \int_k^A \frac{r_n(t)\, dt}{(t - z)^{\alpha+1}} \right| + \left| \int_\delta \frac{r_n(t)\, dt}{(t - z)^{\alpha+1}} \right| \right\}. \tag{3.13}$$

We need the following lemma due to Walsh (op. cit., p. 258).

LEMMA 3.2. *Let C be an arbitrary Jordan curve which contains the origin in its interior. If $| r_n(z) | \leqq M$ on C, then $| r_n(z) | \leqq M\rho^n$, $\rho > 1$, z on E_ρ, and $| r_n(z) | \leqq M/R^n$, $R < 1$, z on I_R.*

Now let C be a curve of type S and let the radius of δ be $M_3(\rho - 1)$, where M_3 is the smaller of M_1 and M_2 in inequalities (2.11). Of course, ρ should be chosen sufficiently near 1 to insure that δ contain neither k nor the origin in its closed interior. Letting $R = 2 - \rho$, we have

$$d(C, E_\rho) \geqq M_1(\rho - 1) \geqq M_3(\rho - 1),$$
$$d(C, I_R) \geqq M_2(1 - R) \geqq M_3(\rho - 1),$$

and consequently δ lies in the closed limited region bounded by E_R and I_ρ. Also,

$$\frac{M}{R^n} = \frac{M}{(2 - \rho)^n} > M\rho^n, \qquad \rho > 1.$$

Hence $| r_n(t) | \leqq M(2 - \rho)^{-n}$ for t on δ. If we take

$$\rho = \frac{n + 2\alpha}{n + \alpha}, \tag{3.14}$$

it is easy to show (see SIII) that

$$| r_n^\alpha(z) | \leqq MK(\alpha, C)n^\alpha, \qquad z \text{ on } C^m, \tag{3.15}$$

for n sufficiently large. Since $r_{n-1}(z)$ is also an $r_n(z)$ with the leading coefficients zero, we can adjust $K(\alpha, C)$ so that (3.15) holds for all n. Thus, we have the following

THEOREM 3.3. *Let C be a curve of type S which contains the origin in its interior. Let $| r_n(z) | \leqq M$ on C. Then we have*

$$| r_n^\alpha(z) | \leqq MK(\alpha, C)n^\alpha, \qquad z \text{ on } C^m,\ \alpha > 0,$$

where $K(\alpha, C)$ is a constant depending only on α and C.

For curves of type u we apply Theorem 2.5. This amounts to replacing α by αu. Thus we have

THEOREM 3.4. *Let C be a curve of type u which contains the origin in its interior. Let $|r_n(z)| \leqq M$ on C. Then we have*

$$|r_n^\alpha(z)| \leqq MK(\alpha, C)n^{\alpha u}, \quad z \text{ on } C^m, \alpha > 0, 1 < u < 2,$$

where $K(\alpha, C)$ is a constant depending only on α and C.

For curves of type W' we apply Theorem 2.6 and we have

THEOREM 3.5. *Let C be a curve of type W' which contains the origin in its interior. Let $|r_n(z)| \leqq M$ on C. Then we have*

$$|r_n^\alpha(z)| \leqq MK(\alpha, C)n^{2\alpha}, \qquad z \text{ on } C^m, \alpha > 0,$$

where $K(\alpha, C)$ is a constant depending only on α and C.

If α is an integer, we may use the Cauchy integral formula for the derivative, the integral being taken around δ. Thus we do not have to integrate along the curve, and hence we have (see SII)

THEOREM 3.6. *Let C be an arbitrary Jordan curve containing the origin in its interior. Let $|r_n(z)| \leqq M$ on C. Then we have*

$$|r_n'(z)| \leqq MK(C)n^2, \qquad z \text{ on } C,$$

where $K(C)$ is a constant depending only on C.

In fact, this theorem extends to an arbitrary ring-shaped region which does not contain the origin in its closed interior and which separates the origin from the point at ∞.

4. **Degree of approximation—analytic curves.** We consider first the case of the unit circle $C: |z| = 1$. Suppose $f(z)$ is defined on C and[11] $f^{(p)}(z)$ satisfies a Lipschitz condition of order β, $0 < \beta \leqq 1$, on C. On C we have $z = e^{i\theta}$ and hence $f(z) \equiv f(e^{i\theta}) \equiv F(\theta)$ and $F^{(p)}(\theta)$ satisfies a Lipschitz condition of order β. Consequently, by Jackson (op. cit., p. 10), we know that there exists a trigonometric polynomial $T_n(\theta)$ of order n in θ such that

$$|F(\theta) - T_n(\theta)| \leqq \frac{M}{n^{p+\beta}},$$

where M is a constant independent of n and θ. But

$$\sin k\theta = \frac{z^k - z^{-k}}{2i}, \qquad \cos k\theta = \frac{z^k + z^{-k}}{2}, \qquad (k = 0, 1, 2, \cdots),$$

and hence $T_n(\theta)$ can be written as a polynomial in z and $1/z$. Thus we have the following

[11] $f^{(p)}(z)$, where p is a positive integer, denotes the p-th derivative of $f(z)$; $f^{(0)}(z) \equiv f(z)$.

THEOREM 4.1. *Let $f^{(p)}(z)$, $p \geqq 0$, satisfy a Lipschitz condition of order β, $0 < \beta \leqq 1$, on $C: |z| = 1$. Then for each n ($n = 1, 2, \cdots$) there exists a polynomial $r_n(z)$ of degree n in z and $1/z$ such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{p+\beta}}, \qquad |z| = 1,$$

where M is a constant independent of n and z.

Now suppose C is an arbitrary analytic Jordan curve in the z-plane containing the origin in its interior. Let $f(z)$ be defined on C and let $f^{(p)}(z)$ satisfy a Lipschitz condition of order β, $0 < \beta \leqq 1$, on C. By mapping the interior C on $|w| < 1$, we have $f(z) = F(w)$, and by Theorem 4.1 there exists $r_n(w)$ such that $|F(w) - r_n(w)| \leqq M/n^{p+\beta}$, $|w| = 1$, or

$$(4.10) \qquad |f(z) - s_n(\phi(z))| \leqq M/n^{p+\beta}, \qquad z \text{ on } C,\ s_n(\phi(z)) = r_n(w).$$

Since $f(z)$ is bounded on C, by Lemma 3.2 we have $|s_n(\phi(z))| \leqq M_1\rho^n$, $\rho > 1$, for z on E_ρ and z on $I_{1/\rho}$. Furthermore, the function $s_n(\phi(z))$ is analytic in this region and can be split into two components, one analytic in the closed exterior of $I_{1/\rho}$ and the other analytic in the closed interior of E_ρ. We can approximate these two components in the closed exterior of $I_{2/(\rho+1)}$ and in the closed interior of $E_{(\rho+1)/2}$ by polynomials $r_n(z)$ in z and $1/z$ so that

$$(4.11) \qquad |s_n(\phi(z)) - r_n(z)| \leqq \frac{M_3}{\sigma^n},$$

for some fixed $\sigma > 1$, where M_3 is independent of σ, n, and z. Thus by combining inequalities (4.10) and (4.11) we have

THEOREM 4.2. *Let C be an analytic Jordan curve containing the origin in its interior. Let $f(z)$ be defined on C and let $f^{(p)}(z)$ satisfy a Lipschitz condition of order β, $0 < \beta \leqq 1$, on C. Then for each n ($n = 1, 2, \cdots$) there exists a polynomial $r_n(z)$ of degree n in z and $1/z$ such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{p+\beta}}, \qquad z \text{ on } C,$$

where M is a constant independent of n and z.

An approximate converse is

THEOREM 4.3. *Let C be an analytic Jordan curve containing the origin in its interior. Let $f(z)$ be defined on C and for each n ($n = 1, 2, \cdots$) let there exist a polynomial $r_n(z)$ of degree n in z and $1/z$ such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{p+\beta}}, \qquad z \text{ on } C,\ 0 < \beta \leqq 1,$$

where M is a constant independent of n and z, and p is a positive integer or zero. Then $f^{(p)}(z)$ exists on C and satisfies the condition

$$|f^{(p)}(z_1) - f^{(p)}(z_2)| \leqq L\,|z_1 - z_2|^\beta\,|\log|z_1 - z_2||^\delta, \qquad z_1, z_2 \text{ on } C,$$

where $\delta = 0$ *if* $\beta < 1$, *and* $\delta = 1$ *if* $\beta = 1$, *and where* L *is a constant independent of* z_1 *and* z_2.

The proof follows a method used by Walsh and Sewell (loc. cit.). For C the unit circle, $f(z)$ is a function of θ and $r_n(z)$ is a trigonometric polynomial of order n in θ. Thus, the results of de la Vallée Poussin (op. cit., Chapter 4) apply directly. For the general analytic Jordan curve the result follows by mapping the interior of C on $|w| < 1$ and approximating $r_n(\Phi(w))$ by a polynomial of degree n in w and $1/w$ (see the proof of Theorem 4.2, and Walsh and Sewell, loc. cit.).

It should be noted that Theorem 4.3 is an exact converse of Theorem 4.2 for $0 < \beta < 1$. An exact converse for $\beta = 1$ is impossible as shown by an example in Walsh and Sewell (loc. cit.).

It has been shown (see SIII) that, if $f^{(p)}(z)$ satisfies a Lipschitz condition of order β on C, $f(z)$ has a bounded derivative of every order $\alpha' < p + \beta$, and that, if $f(z)$ has a bounded derivative of order $\alpha = p + \beta$ on C^m, $f^{(p)}(z)$ satisfies a Lipschitz condition of order β, $0 < \beta \leqq 1$. Of course, if we take two points k, it is clear that the Lipschitz condition holds uniformly on C. Thus in Theorems 4.2 and 4.3 bounded derivatives may be used to describe the continuity properties of the function.

5. **Degree of approximation—more general curves.** Our first result is

THEOREM 5.1. *Let* C *be a curve of type* S *containing the origin in its interior and let* $f(z)$ *be defined on* C. *If for every* n ($n = 1, 2, \cdots$) *there exists a polynomial* $r_n(z)$ *of degree* n *in* z *and* $1/z$ *such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^\alpha}, \qquad \alpha > 0, z \text{ on } C,$$

M *a constant independent of* n *and* z, *then* $f(z)$ *has a bounded derivative of every order* $\alpha' < \alpha$ *on* C.

The proof is based on an application of Theorem 3.3. Using Theorem 16.1 of SIII we can apply a procedure similar to that of Montel[12] to prove the theorem. Of course, this involves two separate choices of k, each yielding a bounded derivative on C^m, and since these two overlap, we have the continuity property for the entire curve.

By applying Theorems 3.4 and 3.5, respectively, and the above method, we obtain

THEOREM 5.2. *Let* C *be a curve of type* u *containing the origin in its interior and let* $f(z)$ *be defined on* C. *If for every* n ($n = 1, 2, \cdots$) *there exists a polynomial* $r_n(z)$ *of degree* n *in* z *and* $1/z$ *such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{\alpha u}}, \qquad \alpha > 0, 1 < u < 2, z \text{ on } C,$$

[12] P. Montel, *Sur les polynomes d'approximation*, Bulletin de la Société Mathématique de France, vol. 46(1919), pp. 151–196.

M a constant independent of n and z, then f(z) has a bounded derivative of every order $\alpha' < \alpha$ on C.

THEOREM 5.3. *Let C be a curve of type W' containing the origin in its interior and let f(z) be defined on C. If for every n ($n = 1, 2, \cdots$) there exists a polynomial $r_n(z)$ of degree n in z and 1/z such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{2\alpha}}, \qquad \alpha > 0, z \text{ on } C,$$

M a constant independent of n and z, then f(z) has a bounded derivative of every order $\alpha' < \alpha$ on C.

For an arbitrary Jordan curve C containing the origin in its interior we apply Theorem 3.6 to obtain (see SII)

THEOREM 5.4. *Let C be an arbitrary Jordan curve containing the origin in its interior and let f(z) be defined on C. If for every n ($n = 1, 2, \cdots$) there exists a polynomial $r_n(z)$ of degree n in z and 1/z such that*

$$|f(z) - r_n(z)| \leqq \frac{M}{n^{\alpha}}, \qquad \alpha > 2, z \text{ on } C,$$

α and M independent of n and z, then f(z) has a bounded first derivative on C.

GEORGIA SCHOOL OF TECHNOLOGY.

THE JUMP OF A FUNCTION DETERMINED BY ITS FOURIER COEFFICIENTS

By Otto Szász

1. Let $f(x)$ be integrable L in the interval $(-\pi, \pi)$ and have period 2π, and let its Fourier series be

$$f(x) \sim \frac{1}{2}a_0 + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x),$$

where

$$2c_\nu = a_\nu - ib_\nu = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)e^{-i\nu t}\, dt \qquad (\nu = 0, 1, 2, \cdots). \tag{1}$$

We shall also use

$$\begin{aligned} s_n(x) &= \frac{a_0}{2} + \sum_1^n (a_\nu \cos \nu x + b_\nu \sin \nu x) \equiv \sum_0^n A_\nu(x), \\ \bar{s}_n(x) &= \sum_1^n (b_\nu \cos \nu x - a_\nu \sin \nu x) \equiv \sum_1^n \bar{A}_\nu(x). \end{aligned} \tag{2}$$

$\bar{s}_n(x)$ is the trigonometric polynomial conjugate to $s_n(x)$. It is well known that the arithmetic means

$$\begin{aligned} \sigma_n(x) &= \frac{s_0 + s_1 + \cdots + s_{n-1}}{n} = \frac{1}{n}\sum_0^n (n - \nu)A_\nu \\ &= \frac{1}{2n\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{\sin \frac{1}{2}n(t - x)}{\sin \frac{1}{2}(t - x)}\right)^2 dt \end{aligned} \tag{3}$$

converge to $\frac{1}{2}\{f(x + 0) + f(x - 0)\}$, whenever this expression exists.[1] We are concerned here with the determination of the jump: $D(x) = f(x + 0) - f(x - 0)$.

Received January 28, 1938; presented to the American Mathematical Society, February 26, 1938.

[1] This is Fejér's classical theorem; it was generalized by Lebesgue, who proved $\sigma_n(x) \to f(x)$ whenever

$$\frac{1}{h}\int_0^h |f(x + t) + f(x - t) - 2f(x)|\, dt \to 0 \qquad \text{as } h \to 0.$$

An extension of this result is [3, §5; 6; 4]:

$$\int_0^h |\varphi(t)|\, dt = O(h) \quad \text{and} \quad \int_0^h \varphi(t)\, dt = o(h),$$

where $\varphi(t) = f(x + t) + f(x - t) - 2f(x)$, imply $\sigma_n(x) \to f(x)$.

Several solutions of this problem have been given. To point out some of them Fejér [2][2] and Csillag [1] proved the following

THEOREM 1. *If $f(x)$ is of bounded variation, then*[3]

$$\pi \lim_{n\to\infty} \{\bar{s}_n(x) - \bar{\sigma}_n(x)\} = f(x+0) - f(x-0), \tag{4}$$

where

$$\bar{\sigma}_n(x) = \frac{\bar{s}_1 + \cdots + \bar{s}_{n-1}}{n}. \tag{5}$$

Lukács [5] proved

THEOREM 2. *Let $f(x)$ be integrable in $(-\pi, \pi)$. If there exists a $D(x)$ such that*

$$\lim_{h\to+0} \frac{1}{h}\int_0^h |f(x+t) - f(x-t) - D(x)|\, dt = 0,$$

then

$$\lim_{n\to\infty} \frac{\bar{s}_n(x)}{\log n} = \frac{1}{\pi} D(x). \tag{6}$$

Finally, the author [7, §3] has given the following

THEOREM 3. *If $f(x)$ is integrable and if there exists a $D(x)$ such that*

$$\lim_{h\to+0} \frac{1}{h}\int_0^h \{f(x+t) - f(x-t) - D(x)\}\, dt = 0, \tag{7}$$

then

$$\lim_{r\to1-0} (1-r)\frac{d}{dx} H(r,x) = \frac{1}{\pi} D(x),$$

where

$$H(r,x) = \frac{a_0}{2} + \sum_1^\infty (a_\nu \cos \nu x + b_\nu \sin \nu x) r^\nu.$$

Remark. We have

$$f(x+t) - f(x-t) \sim 2\sum_1^\infty \bar{A}_\nu(x) \sin \nu t,$$

hence

$$\frac{1}{h}\int_0^h \{f(x+t) - f(x-t)\}\, dt = \pi\left\{\frac{h}{\pi}\sum_1^\infty \nu \bar{A}_\nu(x)\left(\frac{\sin \frac{1}{2}\nu h}{\frac{1}{2}\nu h}\right)^2\right\}.$$

Thus the assumption (7) can be written as

$$\lim_{h\to+0} \frac{2h}{\pi}\sum_1^\infty \nu \bar{A}_\nu(x)\left(\frac{\sin \nu h}{\nu h}\right)^2 = \frac{1}{\pi} D(x);$$

[2] The numbers in brackets refer to the literature at the end of this paper.

[3] Fejér considered functions satisfying Dirichlet's conditions, and Csillag generalized his result to functions of bounded variation. In view of the Riemann-Lebesgue lemma we may replace $\bar{s}_n$ by $\bar{s}_{n+\nu}$, where ν is any fixed integer.

this means that the sequence $\{\nu\bar{A}_\nu\}$ is summable by the Riemann method of the second kind to the value $\pi^{-1}D$. Note that

$$\frac{2h}{\pi}\left\{\frac{1}{2}+\sum_1^\infty\left(\frac{\sin \nu h}{\nu h}\right)^2\right\} \equiv 1.$$

On the other hand, the conclusion of Theorem 3 asserts that the sequence $\{\nu\bar{A}_\nu\}$ is Abel-summable.

2. The purpose of this paper is to give a new determination for $D(x)$ by using a trigonometric polynomial related to the one in (4). We shall first prove

THEOREM I. *Under the same assumption as in Theorem* 2

$$\lim_{n\to\infty}(\bar{\sigma}_{2n}(x)-\bar{\sigma}_n(x)) = \frac{1}{\pi}\log 2\cdot D(x). \tag{8}$$

Remark. Note that (4) can be written as

$$\lim_{n\to\infty} n(\bar{\sigma}_n-\bar{\sigma}_{n-1}) = \lim_{n\to\infty}\frac{1}{n}\sum_1^n \nu\bar{A}_\nu = \frac{1}{\pi}\{f(x+0)-f(x-0)\}.$$

It is the limitation of the sequence $\{\nu\bar{A}_\nu\}$ by arithmetic means. On the other hand, (6) is

$$\lim_{n\to\infty}\frac{(n+1)\bar{\sigma}_{n+1}-n\bar{\sigma}_n}{\log n} \equiv \lim_{n\to\infty}\frac{1}{\log n}\sum_1^n\frac{1}{\nu}\,\nu\bar{A}_\nu = \frac{1}{\pi}D(x).$$

Here logarithmic means are applied.

To prove Theorem I, we write

$$\bar{\sigma}_{2n}-\bar{\sigma}_n = \frac{1}{2n}\sum_1^{2n}(2n-\nu)\bar{A}_\nu - \frac{1}{n}\sum_1^n(n-\nu)\bar{A}_\nu$$
$$= \frac{1}{2n}\left\{\sum_1^n\nu\bar{A}_\nu+\sum_{n+1}^{2n}(2n-\nu)\bar{A}_\nu\right\},$$

where

$$\bar{A}_\nu \equiv \bar{A}_\nu(x) = -\Im(2c_\nu e^{i\nu x}). \tag{9}$$

Now (1) gives

$$2\left\{\sum_1^n \nu c_\nu e^{i\nu x}+\sum_{n+1}^{2n}(2n-\nu)c_\nu e^{i\nu x}\right\}$$
$$=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\left\{\sum_1^n \nu e^{i\nu(x-t)}+\sum_{n+1}^{2n}(2n-\nu)e^{i\nu(x-t)}\right\}dt$$
$$=\frac{e^{ix}}{\pi}\int_{-\pi}^{\pi}e^{-it}f(t)\left\{\frac{1-e^{in(x-t)}}{1-e^{i(x-t)}}\right\}^2 dt = \frac{1}{\pi}\int_{-\pi}^{\pi}e^{in(x-t)}\left(\frac{\sin\frac{1}{2}n(x-t)}{\sin\frac{1}{2}(x-t)}\right)^2 f(t)\,dt.$$

Hence by (9)

$$\bar{\sigma}_{2n} - \bar{\sigma}_n = -\frac{1}{2n\pi}\int_{-\pi}^{\pi} f(t) \sin n(x-t) \left(\frac{\sin \frac{1}{2}n(x-t)}{\sin \frac{1}{2}(x-t)}\right)^2 dt$$

$$= -\frac{1}{2n\pi}\int_{-\pi}^{\pi} f(x-t) \sin nt \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt$$

$$= \frac{1}{2n\pi}\int_{0}^{\pi} \{f(x+t) - f(x-t)\} \sin nt \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt.$$

But $\frac{1}{n\pi}\int_0^\pi \sin nt \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt$ is the n-th arithmetic mean for the cosine series of $\sin nx$ at $x = 0$. On writing

$$\sin nx \sim \frac{\alpha_0}{2} + \sum_1^\infty \alpha_\nu \cos \nu x, \qquad 0 < x < \pi,$$

we have

$$\alpha_\nu = \frac{2}{\pi}\int_0^\pi \sin nt \cos \nu t\, dt = \frac{1}{\pi}\int_0^\pi \{\sin (n+\nu)t + \sin (n-\nu)t\}\, dt$$

$$= \frac{1}{\pi}\frac{2n}{n^2-\nu^2}\{1-(-1)^{n+\nu}\} \qquad (\nu = 0, 1, \cdots, n-1).$$

Hence

$$\frac{1}{n\pi}\int_0^\pi \sin nt \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt = \frac{1}{n\pi}\{1-(-1)^n\} + \frac{1}{n}\sum_1^n (n-\nu)\alpha_\nu$$

$$= \frac{1}{n\pi}\{1-(-1)^n\} + \frac{2}{\pi}\sum_1^n \frac{1-(-1)^{n+\nu}}{n+\nu}.$$

Denoting this expression by ω_n, we next prove that

$$\lim_{n\to\infty} \omega_n = \frac{2}{\pi}\log 2. \tag{10}$$

We have

$$\omega_{2n} = \frac{2}{\pi}\sum_1^{2n} \frac{1-(-1)^\nu}{2n+\nu} = \frac{4}{\pi}\sum_{k=1}^n \frac{1}{2n+2k-1} = \frac{2}{\pi n}\sum_{k=1}^n \frac{1}{1+\dfrac{2k-1}{2n}}.$$

This gives

$$\lim_{n\to\infty} \omega_{2n} = \frac{2}{\pi}\int_0^1 \frac{dx}{1+x} = \frac{2}{\pi}\log 2.$$

Also

$$\omega_{2n+1} = \frac{2}{(2n+1)\pi} + \frac{2}{\pi}\sum_1^{2n}\frac{1-(-1)^{\nu+1}}{2n+\nu+1} = \frac{2}{(2n+1)\pi} + \frac{4}{\pi}\sum_2^{n+1}\frac{1}{2n+2k-1}$$

$$= \omega_{2n} - \frac{2}{(2n+1)\pi} + \frac{4}{(4n+1)\pi},$$

and hence

$$\lim_{n\to\infty} \omega_{2n+1} = \frac{2}{\pi} \log 2.$$

This proves (10). We now get

$$\bar{\sigma}_{2n}(x) - \bar{\sigma}_n(x) - \tfrac{1}{2}\omega_n D(x)$$
$$= \frac{1}{2n\pi} \int_0^\pi \{f(x + t) - f(x - t) - D(x)\} \sin nt \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt,$$

from which it follows that

$$| \bar{\sigma}_{2n}(x) - \bar{\sigma}_n(x) - \tfrac{1}{2}\omega_n D(x) |$$
$$\leqq \frac{1}{2n\pi} \int_0^\pi |f(x + t) - f(x - t) - D(x)| \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt.$$

But it follows from the theorem of Lebesgue already referred to in footnote 1 that for a function $\psi(t)$ for which the mean integral

$$\frac{1}{h} \int_0^h | \psi(t) | \, dt \to 0, \qquad \text{as } h \to 0,$$

we may conclude that

$$\frac{1}{2n\pi} \int_0^\pi | \psi(t) | \left(\frac{\sin \frac{1}{2}nt}{\sin \frac{1}{2}t}\right)^2 dt \to 0, \qquad \text{as } n \to \infty.$$

We finally obtain

$$\lim_{n\to\infty} \{\bar{\sigma}_{2n}(x) - \bar{\sigma}_n(x) - \tfrac{1}{2}\omega_n D(x)\} = 0,$$

or, using (10), we get

$$\lim_{n\to\infty} \{\bar{\sigma}_{2n}(x) - \bar{\sigma}_n(x)\} = \frac{1}{\pi} \log 2 \cdot D(x).$$

This proves the theorem.

Substituting in (8) $n = 2^\nu$, $\nu = 1, 2, 3, \cdots$, and taking the arithmetic means, we find

$$\lim_{n\to\infty} \frac{\bar{\sigma}_{2^n}(x)}{n} = \frac{1}{\pi} \log 2 \cdot D(x).$$

This can be deduced directly from (6).

3. By analogy with the extension of Lebesgue's theorem, referred to in footnote 1, we can extend Theorem I as follows:

Theorem II. *Suppose that $f(x)$ is integrable L in $(-\pi, \pi)$ and that*

$$\int_0^h | \psi(x, t) | \, dt = O(h), \qquad \int_0^h \psi(x, t) \, dt = o(h), \qquad h \to +0,$$

where $\psi(x, t) = f(x + t) - f(x - t) - D(x)$; *then*

$$\lim_{n\to\infty} \{\bar\sigma_{2n}(x) - \bar\sigma_n(x)\} = \frac{1}{\pi}\log 2\cdot D(x).$$

We first prove the

Lemma. $0 < t^2 - \sin^2 t < t^4$ *for* $0 < t \leqq \frac{1}{2}\pi$.

We have for $0 < t < \frac{1}{2}\pi$

$$0 < t < \frac{\sin t}{\cos t},$$

hence

$$t^2 < \frac{\sin^2 t}{1 - \sin^2 t},$$

or

$$t^2 < (1 + t^2)\sin^2 t, \qquad \sin^2 t > \frac{t^2}{1+t^2} > t^2 - t^4.$$

Finally

$$0 < t^2 - \sin^2 t < t^4.$$

We now write

$$\begin{aligned}\bar\sigma_{2n} - \bar\sigma_n - \tfrac{1}{2}\omega_n D &= \frac{1}{2n\pi}\int_0^\pi \psi(x, t)\,\frac{\sin^2 \frac{1}{2}nt}{(\frac{1}{2}t)^2}\sin nt\,dt \\ &\quad - \frac{1}{2n\pi}\int_0^\pi \psi(x,t)\sin^2\tfrac{1}{2}nt\cdot\sin nt\left\{\frac{1}{(\frac{1}{2}t)^2} - \frac{1}{\sin^2\frac{1}{2}t}\right\}dt \\ &\equiv I_1^{(n)} + I_2^{(n)}.\end{aligned}$$

But by the lemma

$$|I_2^{(n)}| \leqq \frac{1}{2n\pi}\int_0^\pi |\psi(x,t)|\left(\frac{t}{2}\right)^4\left(\frac{t}{2}\right)^{-2}\frac{dt}{\sin^2\frac{1}{2}t} \leqq \frac{\pi}{2n}\cdot\frac{1}{4}\int_0^\pi |\psi(x,t)|\,dt = O\left(\frac{1}{n}\right).$$

We thus have only to prove $I_1^{(n)} \to 0$. On putting

$$\max_{0<h\leqq\delta}\frac{1}{h}\int_0^h \psi(x,t)\,dt = \epsilon(\delta),$$

we have $\epsilon(\delta) \downarrow 0$ as $\delta \downarrow 0$. Thus $\lambda(\delta) \equiv \delta^{-1}\epsilon^{-\frac{1}{2}} \uparrow \infty$ as $\delta \downarrow 0$, so that if $\delta = \rho(\lambda)$ is the function inverse to λ, then $\rho(\lambda) \downarrow 0$ as $\lambda \uparrow \infty$, and $\lambda\cdot\rho(\lambda) \equiv \delta^{-1}\epsilon^{-\frac{1}{2}}\delta \equiv \epsilon^{-\frac{1}{2}} \uparrow \infty$ as $\delta \downarrow 0$, or as $\lambda \uparrow \infty$. We now split up the expression $I_1^{(n)}$ in the form

$$I_1^{(n)} = \frac{1}{2n\pi}\left\{\int_0^{\rho(n)} + \int_{\rho(n)}^\pi\right\} \equiv T_1^{(n)} + T_2^{(n)}, \qquad n > n_0,$$

where n_0 is chosen so large that $\rho(n_0) < \pi$. Writing

$$\int_0^h \psi(x, t)\,dt = \psi_1(h), \qquad \int_0^h |\psi(x, t)|\,dt = \eta(h),$$

we find

$$|T_2^{(n)}| \leqq \frac{1}{2n\pi}\int_{\rho(n)}^{\pi} |\psi(x, t)|\frac{4}{t^2}\,dt = \frac{2}{n\pi}\left\{t^{-2}\eta(t)\Big|_{\rho(n)}^{\pi} + 2\int_{\rho(n)}^{\pi} t^{-3}\eta(t)\,dt\right\}.$$

Hence

$$|T_2^{(n)}| < \frac{2}{n\pi}\left\{\frac{\eta(\pi)}{\pi^2} + 2O\left(\int_{\rho(n)}^{\pi} t^{-2}\,dt\right)\right\} = O\left(\frac{1}{n} + \frac{1}{n\rho(n)}\right) \to 0 \quad \text{as} \quad n \to \infty.$$

Furthermore

$$T_1^{(n)} = \frac{4}{n\pi}\int_0^{\rho(n)} \psi(x, t)\,\frac{(1 - \cos nt)\sin nt}{t^2}\,dt$$

$$= \frac{4}{n\pi}\int_0^{\rho(n)} \psi(x, t)\,\frac{\sin nt - \frac{1}{2}\sin 2nt}{t^2}\,dt$$

$$= \frac{4}{n\pi}\left\{\psi_1(t)\,\frac{\sin nt - \frac{1}{2}\sin 2nt}{t^2}\Big|_0^{\rho(n)} - \int_0^{\rho(n)} \psi_1(t)\,\frac{n(\cos nt - \cos 2nt)}{t^2}\,dt - \int_0^{\rho(n)} \psi_1(t)\,\frac{2(\sin nt - \frac{1}{2}\sin 2nt)}{t^3}\,dt\right\}.$$

Hence, writing ρ_n for $\rho(n)$, we finally obtain

$$|T_1^{(n)}| < \frac{4}{\pi}2\epsilon(\rho_n) + \frac{8}{\pi}\epsilon(\rho_n)\int_0^{\rho_n} t^{-1}\sin^2 nt\,dt + \frac{8}{n\pi}\epsilon(\rho_n)\int_0^{\rho_n} 2t^{-2}|\sin nt|\sin^2 \tfrac{1}{2}n\,dt$$

$$< \frac{8}{\pi}\left\{\epsilon(\rho_n) + \epsilon(\rho_n)n^2\int^{\rho_n} t\left(\frac{\sin nt}{nt}\right)^2 dt + \frac{1}{n}\epsilon(\rho_n)\int_0^{\rho_n}\left(\frac{\sin \frac{1}{2}nt}{\frac{1}{2}nt}\right)^2 \tfrac{1}{2}n^2\,dt\right\}$$

$$< \frac{8}{\pi}\epsilon(\rho_n)\{1 + \tfrac{1}{2}n^2\rho_n^2 + \tfrac{1}{2}n\rho_n\} = \frac{8}{\pi}\epsilon(\rho_n)\{1 + \tfrac{1}{2}\epsilon(\rho_n)^{-1} + \tfrac{1}{2}\epsilon(\rho_n)^{-\frac{1}{2}}\} \to 0$$

as $n \to \infty$. This proves Theorem II.

REFERENCES

1. P. CSILLAG, *Korlátos ingadozású függvények Fourier-féle állandóiról*, Mathematikai és Physikai Lapok, vol. 27(1918), pp. 301–308.
2. L. FEJÉR, *Über die Bestimmung des Sprunges der Funktion aus ihrer Fourierreihe*, Journal für Mathematik, vol. 142(1913), pp. 165–188.
3. G. H. HARDY AND J. E. LITTLEWOOD, *The Fourier series of a positive function*, Journal of the London Mathematical Society, vol. 1(1926), pp. 134–138.
4. M. JACOB, *Bemerkung zur Konvergenz und Summierbarkeit Fourierscher Reihen*, ibid., vol. 3(1928), pp. 124–129.
5. F. LUKÁCS, *Über die Bestimmung des Sprunges . . .*, Journal für Mathematik, vol. 150 (1920), pp. 107–112.
6. S. POLLARD, *Direct proof of a new criterion for the summability of a Fourier series*, Journal of the London Mathematical Society, vol. 1(1926), pp. 233–235.
7. O. SZÁSZ, *Zur Konvergenztheorie der Fourierschen Reihen*, Acta Mathematica, vol. 61 (1933), pp. 185–201.

UNIVERSITY OF CINCINNATI.

INTEGRATION IN ABSTRACT METRIC SPACES

By S. Saks

1. In a recent note Banach[1] established the following theorem which may be regarded as an interesting extension of the well-known formula of F. Riesz for linear functionals over the space of continuous functions on a finite interval.

Let **H** be a compact metric space and Φ a non-negative linear functional defined over the space of continuous real functions on **H**; i.e.,

(i) $\Phi(f) \geqq 0$ whenever f is a non-negative continuous function on **H**,

(ii) $\Phi(f + g) = \Phi(f) + \Phi(g)$ for any two continuous functions f and g on **H**,

(iii) $\lim_n \Phi(f_n) = 0$, if $\{f_n\}$ is a sequence of continuous functions, converging to 0 uniformly on **H**.

Then there exists a measure μ in the space **H**, with respect to which

$$\Phi(g) = \int_{\mathbf{H}} g(x)\, d\mu(x) \tag{1.1}$$

for any function g continuous on **H**.[2]

Banach's original proof of the above theorem is based on the general theory of functional operations and on his theory of integration on abstract spaces. In this note we give another proof which seems more elementary and which is based directly on the Lebesgue theory of integration as extended to abstract spaces by Fréchet.

2. In what follows **H** will be a fixed compact metric space and $\rho(a, b)$ the distance between any two points a, b of the space. If $a \,\varepsilon\, \mathbf{H}$ and $r > 0$, then $S(a, r)$ will denote the open sphere with center a and radius r, i.e., the set of points x such that $\rho(a, x) < r$. The set of points x such that $\rho(a, x) = r$ is the surface of the sphere $S(a, r)$. If A is any set in **H**, the closure of A will be denoted, as usual, by $\bar{A}$.

Finally, Φ will denote a non-negative linear functional defined over the space of continuous functions on **H**.

3. For every set E in **H**, we shall denote by $\Gamma(E)$ the lower bound of the numbers $\Phi(f)$, where f is an arbitrary non-negative continuous function on **H**

Received February 20, 1938.

[1] S. Banach, *The Lebesgue integral in abstract spaces* (Note II in the book by S. Saks, *Theory of the Integral*, 2d ed., Monografie Matematyczne, Warsawa, 1937, pp. 320–330, esp. p. 326).

[2] In this connection see also G. Fichtenholz and L. Kantorovich, *Sur les opérations dans l'espace des fonctions bornées*, Studia Mathematica, vol. 5(1934), pp. 67–78.

such that $f(x) \geqq 1$ for $x \in E$. Thus defined, the set function Γ is clearly non-negative and satisfies the following three conditions:

(P_1) $\Gamma(A) \leqq \Gamma(B)$ whenever $A \subseteq B$,

(P_2) $\Gamma(A + B) \leqq \Gamma(A) + \Gamma(B)$ for any pair of sets A, B in $\mathbf{H}$,

(P_3) $\Gamma(A + B) = \Gamma(A) + \Gamma(B)$ whenever $\rho(A, B) > 0$,

where $\rho(A, B)$ is the distance between the sets A and B.

Only the property (P_3) requires a proof, the first two properties being obvious. Since $\rho(A, B) > 0$, by a well-known theorem[3] there exists a function $h(x)$ continuous in the whole space $\mathbf{H}$, and equal to 1 on B and to 0 on A.

Now let ϵ be an arbitrary positive number and f a non-negative function continuous on $\mathbf{H}$ subject to the conditions

$$f(x) \geqq 1 \qquad \text{for } x \in A + B, \tag{3.1}$$

$$\Gamma(A + B) + \epsilon \geqq \Phi(f). \tag{3.2}$$

Put $f_1 = (1 - h)f$, $f_2 = hf$. Both functions f_1 and f_2 are non-negative and continuous, and by (3.1), we have

$$f_1(x) = f(x) \geqq 1 \text{ on } A, \qquad f_2(x) = f(x) \geqq 1 \text{ on } B.$$

Consequently, in virtue of (3.2),

$$\Gamma(A + B) + \epsilon \geqq \Phi(f) = \Phi(f_1) + \Phi(f_2) \geqq \Gamma(A) + \Gamma(B),$$

whence $\Gamma(A + B) \geqq \Gamma(A) + \Gamma(B)$, and because of the property (P_2), $\Gamma(A + B) = \Gamma(A) + \Gamma(B)$.

4. For any set E in $\mathbf{H}$ we shall denote by $\mu(E)$ the lower bound of the sums $\sum_k \Gamma(G_k)$, where $\{G_k\}$ is an arbitrary sequence of open sets such that $E \subseteq \sum_k G_k$. From the properties (P_1), (P_2) and (P_3), it immediately follows that the function of sets μ satisfies the three conditions of outer measure of Carathéodory:[4]

(C_1) $\mu(A) \leqq \mu(B)$ whenever $A \subseteq B$,

(C_2) $\mu(\sum_k A_k) \leqq \sum_k \mu(A_k)$ for any finite or infinite sequence $\{A_k\}$ of sets,

(C_3) $\mu(A + B) = \mu(A) + \mu(B)$ whenever $\rho(A, B) > 0$.

Furthermore, if F is any closed set, then $\mu(F) = \Gamma(F)$. In fact, if $\{G_k\}$ is a sequence of open sets such that $F \subset \sum_k G_k$ then, by the Borel covering theorem, for sufficiently large N we have $F \subset \sum_{k=1}^{N} G_k$ and consequently, in view of the properties (P_1), (P_2),

$$\Gamma(F) \leqq \sum_{k=1}^{N} \Gamma(G_k) \leqq \sum_{k=1}^{\infty} \Gamma(G_k).$$

[3] See, e.g., P. Alexandroff and H. Hopf, *Topologie*, vol. I, p. 74.

[4] See, e.g., S. Saks, loc. cit., p. 43.

Since $\{G_k\}$ is an arbitrary sequence of open sets covering F, it follows that $\Gamma(F) \leqq \mu(F)$.

To establish the opposite inequality, let ϵ be an arbitrary positive number and f a non-negative function continuous on $\mathbf{H}$ such that $f(x) \geqq 1$ for $x \in F$ and such that $\Phi(f) \leqq \Gamma(F) + \epsilon$. Let G be the set of points x at which $(1 + \epsilon)f(x) > 1$. The set G clearly is open and contains the set F. Hence,

$$\mu(F) \leqq \Gamma(G) \leqq \Phi[(1 + \epsilon)f] = (1 + \epsilon)\Phi(f) \leqq (1 + \epsilon)[\Gamma(F) + \epsilon].$$

Therefore $\mu(F) \leqq \Gamma(F)$. This completes the proof.[5]

5. The function of sets μ, as an outer measure in the sense of Carathéodory, determines a class of measurable sets and a process of Lebesgue integration. In order to establish the theorem of Banach stated in §1, we only have to prove that the inequality

$$\Phi(g) \leqq \int_{\mathbf{H}} g(x)\, d\mu(x) \tag{5.1}$$

holds for any continuous function g on $\mathbf{H}$. The opposite inequality will then immediately follow by changing the sign of g. On the other hand, by adding a constant to g, if necessary, we can confine ourselves to the case when g is continuous and non-negative.

Now let ϵ be an arbitrary positive number, and η a positive number such that $|g(x_2) - g(x_1)| < \epsilon$ whenever $\rho(x_1, x_2) < \eta$. By the Borel-Lebesgue covering theorem the space $\mathbf{H}$ may be covered by a finite number of open spheres $S_1, S_2, \cdots, S_n$ with radii less than $\frac{1}{2}\eta$. Moreover, since $\mu(\mathbf{H}) < \infty$, we may assume that the surface of each of these spheres is of measure (μ) zero.[6]

Now let $K_1 = \bar{S}_1$, $K_2 = \overline{S_2 - K_1}, \cdots, K_n = \overline{S_n - K_{n-1}}$. The sets $K_1, K_2, \cdots, K_n$ form a finite system of closed, non-overlapping sets of diameters less than η. Furthermore, the boundary of each of these sets is of measure (μ) zero; hence on denoting by l_i the lower bound of g on K_i $(i = 1, 2, \cdots, n)$, we have

$$\int_{\mathbf{H}} g(x)\, d\mu(x) \geqq \sum_{i=1}^{n} l_i \mu(K_i). \tag{5.2}$$

On the other hand, since $\mu(K_i) = \Gamma(K_i)$ $(i = 1, 2, \cdots, n)$ (cf. §4), we can associate with each K_i a non-negative continuous function f_i such that $f_i(x) \geqq 1$ on K_i and such that $\mu(K_i) \geqq \Phi(f_i) + \epsilon/(nl_i)$. Put $f(x) = l_1 f_1(x) + \cdots + l_n f_n(x) + \epsilon$. Since the oscillation of g on each set K_i does not exceed ϵ, it

[5] The above proof shows that the inequality $\mu(F) \leqq \Gamma(F)$ holds for any set F whatsoever, not necessarily a closed set.

[6] Indeed, if a is any fixed point, the set of values $r \geqq 0$ for which the surface of the sphere $S(a, r)$ is of positive measure (μ) is at most enumerable.

follows that $f(x) \geqq l_i f_i(x) + \epsilon \geqq g(x)$ for $x \in K_i$ $(i = 1, 2, \cdots, n)$, and so $f(x) \geqq g(x)$ on the whole space $\mathbf{H}$. Thus, in virtue of (5.2),

$$\int_{\mathbf{H}} g(x)\, d\mu(x) \geqq \sum_{i=1}^{n} l_i[\Phi(f_i) + \epsilon/(nl_i)] = \Phi\left(\sum_{i=1}^{n} l_i f_i\right) + \epsilon$$
$$= \Phi(f) - \Phi(\epsilon) + \epsilon \geqq \Phi(g) - \Phi(\epsilon) + \epsilon.$$

Allowing here ϵ to approach 0, we obtain the inequality (5.1), and the proof is complete.

6. The formula (1.1) holds for all linear functionals Φ, not, however, necessarily non-negative, under the condition that μ is interpreted as a general completely additive function of Borel sets. In fact, any linear (i.e., additive and continuous) functional over the space of continuous functions on $\mathbf{H}$ may be represented as the difference of two non-negative linear functionals. We thus obtain a complete generalization of F. Riesz' formula mentioned at the beginning of this note.

UNIVERSITY OF WARSAW.

THE PROJECTIONS OF THE ASYMPTOTIC CURVES

By M. L. MacQueen

1. **Introduction.** A line l_1 through a point of a surface in ordinary space but not lying in the tangent plane of the surface at the point and a line l_2 lying in the tangent plane but not passing through the point are called reciprocal lines if they are reciprocal polars with respect to the quadric of Lie at the point.

G. M. Green,[1] in his investigation of the theory of reciprocal congruences, arrived at an important pair of reciprocal lines, now commonly called the canonical edges of Green, by considering the projections of the asymptotic curves upon the tangent plane at a point of a surface.

In this paper we propose to continue the investigation of the projections of the asymptotic curves upon the tangent plane at a point of a surface. For this purpose power series expansions for the projected asymptotics are deduced, the center of projection being a point on an arbitrary line l_1 at a point of the surface. Consideration of certain osculants associated with the projected asymptotic curves leads to new geometric characterizations of the canonical edges of Green and to other canonical lines. Finally, brief attention is given to a particular transformation of Čech.

2. **The projections of the asymptotic curves.** If the four homogeneous projective coördinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x on a non-ruled surface S in ordinary space are given as analytic functions of two independent variables u, v, and if the parametric net on S is the asymptotic net, then the functions x are solutions of a system of differential equations which, by suitable choice of proportionality factor, can be reduced to Fubini's canonical form

$$(1)\qquad \begin{aligned} x_{uu} &= px + \theta_u x_u + \beta x_v, \\ x_{vv} &= qx + \gamma x_u + \theta_v x_v, \end{aligned} \qquad (\theta = \log \beta\gamma).$$

The coefficients of these equations are functions of u, v and satisfy three integrability conditions.

The coördinates X of a point near P_x and on the u-curve through P_x are given by an expansion of the form

$$(2)\qquad X = x + x_u \Delta u + \tfrac{1}{2} x_{uu} \Delta u^2 + \cdots.$$

Received March 2, 1938.

[1] G. M. Green, *Memoir on the general theory of surfaces and rectilinear congruences*, Transactions of the American Mathematical Society, vol. 20(1919), p. 108.

If the points x, x_u, x_v, x_{uv} are used as the vertices of a local tetrahedron of reference with a suitably chosen unit point, the local coördinates $y_1, \cdots, y_4$ of the point X are represented by the expansions

$$
\begin{aligned}
y_1 &= 1 + \tfrac{1}{2}p\Delta u^2 + \cdots, \\
y_2 &= \Delta u + \tfrac{1}{2}\theta_u \Delta u^2 + \cdots, \\
y_3 &= \tfrac{1}{2}\beta\Delta u^2 + \tfrac{1}{6}(\beta_u + \beta\theta_u)\Delta u^3 + \cdots, \\
y_4 &= \tfrac{1}{6}\beta\Delta u^3 + \tfrac{1}{12}(\beta_u + \beta\theta_u)\Delta u^4 + \cdots.
\end{aligned} \tag{3}
$$

In the notation employed by Lane,[2] a line l_1 through a general point P_x of a surface joins the point x to the point y defined by placing

$$y = -ax_u - bx_v + x_{uv}, \tag{4}$$

wherein a, b are functions of u, v. Dually, a line l_2 in the tangent plane at the point P_x joins the points ρ, σ defined by placing

$$\rho = x_u - bx, \qquad \sigma = x_v - ax, \tag{5}$$

where a, b are functions of u, v. If the functions a, b are the same in equations (4), (5), the lines l_1, l_2 are called *reciprocal lines*, because they are reciprocal polar lines with respect to the quadric of Lie or any quadric of Darboux at the point P_x. Moreover, two reciprocal lines l_1, l_2 are *canonical lines* of the first and second kind respectively in case

$$a = -k\psi, \qquad b = -k\varphi, \tag{6}$$

where k is a constant and

$$\varphi = (\log \beta\gamma^2)_u, \qquad \psi = (\log \beta^2\gamma)_v. \tag{7}$$

If the local coördinates of a point referred to the tetrahedron x, x_u, x_v, x_{uv} are $y_1, \cdots, y_4$, and if the coördinates of the same point are $x_1, \cdots, x_4$ when referred to the tetrahedron x, ρ, σ, $y + \lambda x$, where λ is an arbitrary scalar function of u, v, then the equations of the transformation of coördinates between the two tetrahedrons can be written in the form

$$
\begin{aligned}
x_1 &= y_1 + by_2 + ay_3 + (2ab - \lambda)y_4, \\
x_2 &= y_2 + ay_4, \\
x_3 &= y_3 + by_4, \\
x_4 &= y_4.
\end{aligned} \tag{8}
$$

The parametric equations of the projection of the asymptotic u-curve from the new vertex $(0, 0, 0, 1)$ onto the tangent plane, $x_4 = 0$, are found by substituting the series (3) for $y_1, \cdots, y_4$ into equations (8) and taking such linear combina-

[2] E. P. Lane, *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press, 1932, p. 82.

tions of the resulting coördinates $x_1, \cdots, x_4$ and of 0, 0, 0, 1 as will make the fourth coördinate vanish. For these results we find

$$
(9)\qquad
\begin{aligned}
x_1 &= 1 + b\Delta u + \cdots, \\
x_2 &= \Delta u + \tfrac{1}{2}\theta_u \Delta u^2 + \cdots, \\
x_3 &= \tfrac{1}{2}\beta\Delta u^2 + \tfrac{1}{6}(\beta_u + \beta\theta_u + b\beta)\Delta u^3 + \cdots.
\end{aligned}
$$

Introducing non-homogeneous coördinates in the tangent plane by the definitions

$$
(10)\qquad x = x_2/x_1, \qquad y = x_3/x_1,
$$

we find

$$
(11)\qquad
\begin{aligned}
x &= \Delta u + \tfrac{1}{2}(\theta_u - 2b)\Delta u^2 + \cdots, \\
y &= \tfrac{1}{2}\beta\Delta u^2 + \tfrac{1}{6}(\beta_u + \beta\theta_u - 2b\beta)\Delta u^3 + \cdots.
\end{aligned}
$$

Inverting the first of these series, we obtain

$$
(12)\qquad \Delta u = x - \tfrac{1}{2}(\theta_u - 2b)x^2 + \cdots.
$$

If we substitute this series for Δu in the second of the series (11), we arrive at *the power series expansion for the projection of the asymptotic u-curve from a point on a line l_1 upon the tangent plane*, namely,

$$
(13)\qquad y = \tfrac{1}{2}\beta x^2 - \tfrac{1}{6}\beta(\varphi - 4b)x^3 + \cdots.
$$

Similar calculations lead to the following expansion for the projection of the v-curve upon the tangent plane,

$$
(14)\qquad x = \tfrac{1}{2}\gamma y^2 - \tfrac{1}{6}\gamma(\psi - 4a)y^3 + \cdots.
$$

It will be observed that, as far as written, equations (13), (14) are independent of λ and hence of the position of the center of projection on the line l_1 which is now being used as the edge $x_2 = x_3 = 0$ of the tetrahedron of reference.

3. **The osculating nodal cubic of the projected asymptotics.** The plane cubic curve which has the point P_x for a node and the asymptotic tangents $x_2 = 0$ and $x_3 = 0$ for nodal tangents has the equation

$$
(15)\qquad x_1x_2x_3 + a_1x_2^3 + b_1x_2^2x_3 + b_2x_2x_3^2 + a_2x_3^3 = 0.
$$

If we introduce non-homogeneous coördinates by the definitions (10) and make use of equations (13), (14), we find that the nodal cubic (15) has second order contact with each of the projected asymptotic curves at the point P_x in case

$$
(16)\qquad a_1 = -\tfrac{1}{2}\beta, \qquad a_2 = -\tfrac{1}{2}\gamma.
$$

Moreover, it has third order contact with each of the projections in case

$$(17)\qquad b_1 = \tfrac{1}{3}(\varphi - 4b), \qquad b_2 = \tfrac{1}{3}(\psi - 4a).$$

Therefore, *the equation of the nodal cubic curve having contact of the third order with the projections of the asymptotic curves at the point P_x is given by*

$$(18)\qquad xy + \tfrac{1}{3}(\varphi - 4b)x^2y + \tfrac{1}{3}(\psi - 4a)xy^2 - \tfrac{1}{2}(\beta x^3 + \gamma y^3) = 0.$$

The line containing the three inflexions of the cubic (18) has the equation

$$(19)\qquad x_1 + \tfrac{1}{3}(\varphi - 4b)x_2 + \tfrac{1}{3}(\psi - 4a)x_3 = 0.$$

This line obviously coincides with the line $\rho\sigma$, $x_1 = 0$, in case

$$(20)\qquad a = \tfrac{1}{4}\psi, \qquad b = \tfrac{1}{4}\varphi.$$

Thus we obtain a new geometric characterization of the first canonical edge of Green which may be stated in the following words:

The line of inflexions of the osculating nodal cubic of the projected asymptotic curves coincides with the line $\rho\sigma$ if, and only if, the center of projection is a point on the first canonical edge of Green.

If we suppose that the center of projection is a point on the projective normal at P_x so that $a = b = 0$, we obtain the following result:

If the asymptotic curves are projected onto the tangent plane from a point on the projective normal, the line of inflexions of the osculating nodal cubic of the projected asymptotics is the second axis of Čech.

4. **Conics associated with the projected asymptotics.** By means of the power series expansion (13), we find that the conics having contact of the third order with the projected u-curve at the point P_x are given by

$$(21)\qquad y - \tfrac{1}{2}\beta x^2 + \tfrac{1}{3}(\varphi - 4b)xy + hy^2 = 0,$$

where h is a parameter. The particular conic, K_u, of the pencil (21) which passes through the point σ has the equation

$$(22)\qquad y - \tfrac{1}{2}\beta x^2 + \tfrac{1}{3}(\varphi - 4b)xy = 0.$$

Similarly, the conic K_v having contact of the third order with the projected v-curve at P_x and passing through the point ρ has the equation

$$(23)\qquad x - \tfrac{1}{2}\gamma y^2 + \tfrac{1}{3}(\psi - 4a)xy = 0.$$

The equation of the tangent to the conic K_u at the point σ is found to be

$$(24)\qquad x_1 + \tfrac{1}{3}(\varphi - 4b)x_2 = 0,$$

and the tangent to the conic K_v at the point ρ is given by

$$(25)\qquad x_1 + \tfrac{1}{3}(\psi - 4a)x_3 = 0.$$

These two tangents coincide with the line $\rho\sigma$ if, and only if,

$$a = \tfrac{1}{4}\psi, \qquad b = \tfrac{1}{4}\varphi.$$

Thus the following theorem is proved:

If the asymptotic curves at a point P_x of a surface are projected from any point of a line l_1 onto the tangent plane at the point, and if the four-point conics K_u, K_v at the point P_x of the projected asymptotics pass through the points σ, ρ, respectively, then these conics are tangent to the line l_2 if, and only if, the lines l_1, l_2 are the canonical edges of Green.

Let us suppose that the center of projection is a point on any canonical line of the first kind distinct from the first edge of Green. Use of equations (6) shows that the two tangents defined by equations (24), (25) can be written respectively in the form

$$\begin{aligned} &x_1 + \tfrac{1}{3}(1 + 4k)\varphi x_2 = 0, \\ &x_1 + \tfrac{1}{3}(1 + 4k)\psi x_3 = 0, \end{aligned} \qquad (1 + 4k \neq 0). \tag{26}$$

These two lines are found to intersect in a point which lies on the first canonical tangent at P_x. Thus we arrive at the theorem:

Let the asymptotic curves at a point P_x of a surface be projected upon the tangent plane from a point on any canonical line of the first kind distinct from the first edge of Green. Let the four-point conics K_u, K_v at the point P_x of the projected asymptotics pass through the points σ, ρ, respectively. The lines tangent to these conics at the points ρ, σ intersect in a point which lies on the first canonical tangent at P_x.

It is a routine matter to calculate[3] a power series expansion for one non-homogeneous projective coördinate z of a point on a surface S in terms of the other two coördinates x, y. Referred to the tetrahedron x, x_u, x_v, x_{uv} with suitably chosen unit point, such an expansion takes the form

$$\begin{aligned} z = xy &- \tfrac{1}{3}(\beta x^3 + \gamma y^3) \\ &+ \tfrac{1}{12}(\beta\varphi x^4 - 4\beta\psi x^3 y - 6\theta_{uv} x^2 y^2 - 4\gamma\varphi x y^3 + \gamma\psi y^4) + \cdots. \end{aligned} \tag{27}$$

It is well known that the tangent plane, $z = 0$, at an ordinary point of a surface intersects the surface in a plane curve C with a node at the point, the nodal tangents being the asymptotic tangents of the surface at the point. The two branches of the curve C which are tangent to the u-tangent, $z = y = 0$, and the v-tangent, $z = x = 0$, will be referred to as the u- and v-branches, respectively. The u- and v-branches of the curve C are easily found to be given respectively by the expansions

$$\begin{aligned} y &= \tfrac{1}{3}\beta x^2 - \tfrac{1}{12}\beta\varphi x^3 + \cdots, \\ x &= \tfrac{1}{3}\gamma y^2 - \tfrac{1}{12}\gamma\psi y^3 + \cdots. \end{aligned} \tag{28}$$

[3] E. P. Lane, *Power series expansions in the neighborhood of a point on a surface*, Proceedings of the National Academy of Sciences, vol. 13(1927), p. 808.

Using a power series expansion for a surface which differs from the expansion (27), D. Sun has studied[4] certain osculants of the curve of intersection of a surface and its tangent plane. For example, in our notation, the nodal cubic (15) which has third order contact at the point P_x with both u- and v-branches of the curve C is found to have the equation

$$xy + \tfrac{1}{4}\varphi x^2 y + \tfrac{1}{4}\psi xy^2 - \tfrac{1}{3}(\beta x^3 + \gamma y^3) = 0. \tag{29}$$

The line of inflexions of this cubic[5] is the second canonical edge of Green. Moreover, all of the ∞^4 non-composite cubic surfaces having fourth order contact with the surface at the point P_x cut the tangent plane of the surface at P_x in the cubic (29).

Let C_1 and C_2 be two plane curves having contact of the first order at a point P. Let us consider an arbitrary conic having four-point contact with the curve C_1 at P and another conic having similar contact with C_2 at P. These two conics intersect twice at P and in two other points M, N. Bompiani has shown[6] that the harmonic conjugate of the common tangent of the two curves with respect to the two lines PM, PN is an invariant line which is independent of the two particular conics of the pencils considered.

Since the projected u-curve and the u-branch of the curve of intersection of the surface and the tangent plane have contact of the first order at an ordinary point P_x, we may therefore apply the considerations summarized in the preceding paragraph. For this purpose, using the triangle of reference $x\rho\sigma$, we find that the equation of the four-point conics at the point P_x of the u-branch of the curve of intersection of the surface and its tangent plane at P_x is

$$y - \tfrac{1}{3}\beta x^2 + \tfrac{1}{4}(\varphi - 4b)xy + ky^2 = 0, \tag{30}$$

where k is a parameter. Subtracting this equation from equation (21), we obtain

$$\beta x^2 - \tfrac{1}{2}(\varphi - 4b)xy - 6(h - k)y^2 = 0, \tag{31}$$

which is the equation of the two lines projecting from P_x the points of intersection of the two conics. It is easy to show that the harmonic conjugate of the common u-tangent, $y = 0$, with respect to the two lines given by equation (31) is the line

$$4\beta x - (\varphi - 4b)y = 0. \tag{32}$$

Consideration of the projected v-curve and the v-branch of the curve C at the point P_x yields an analogous line whose equation is

$$(\psi - 4a)x - 4\gamma y = 0. \tag{33}$$

[4] D. Sun, *On the curve of intersection of a surface and its tangent plane*, Tôhoku Mathematical Journal, vol. 38(1933), p. 245.

[5] Fubini and Čech, *Introduction à la géométrie projective différentielle des surfaces*, Paris, 1931, p. 100.

[6] E. Bompiani, *Invarianti proiettivi di contatto fra curve piane*, Rendiconti dei Lincei, ser. 6, vol. 3(1926), p. 118.

The lines (32), (33) coincide respectively with the v-tangent, $x = 0$, and the u-tangent, $y = 0$, in case

$$a = \tfrac{1}{4}\psi, \qquad b = \tfrac{1}{4}\varphi.$$

Consequently we have the following theorem:

Consider an arbitrary conic having four-point contact with the projected u-curve (v-curve) at the point P_x and another conic having similar contact with the u-branch (v-branch) of the curve of intersection of the surface and its tangent plane at P_x. The harmonic conjugate of the common u-tangent (v-tangent) with respect to the two lines projecting from P_x the two points of intersection of these two conics coincides with the v-tangent (u-tangent) if, and only if, the asymptotic curves are projected from a point on the first canonical edge of Green at the point P_x.

5. **The transformation of Čech.** Let us consider the transformation[7]

$$(34)\qquad \begin{gathered} \rho x_1 = -\xi_2\xi_3\xi_4 + k(\beta\xi_3^3 + \gamma\xi_2^3), \\ \rho x_2 = \xi_2\xi_3^2, \qquad \rho x_3 = \xi_2^2\xi_3, \qquad x_4 = 0, \end{gathered}$$

where ρ is a proportionality factor and k is an arbitrary constant. It will be observed that the transformation of Čech (34) is a transformation between the points with local coördinates x in the tangent plane of the surface at a point and planes with local coördinates ξ through the point.

An arbitrary plane

$$(35)\qquad \lambda x_2 + \mu x_3 = 0$$

through a line l_1 at the point P_x of the surface has plane coördinates

$$(36)\qquad (0, \lambda, \mu, 0).$$

The points corresponding to the planes through the line l_1 in the transformation (34) are given by

$$(37)\qquad \begin{gathered} \rho x_1 = k(\beta\mu^3 + \gamma\lambda^3), \\ \rho x_2 = \lambda\mu^2, \qquad \rho x_3 = \lambda^2\mu, \qquad x_4 = 0. \end{gathered}$$

Homogeneous elimination of λ, μ and use of equations (10) yields the equation of a pencil of cubic curves

$$(38)\qquad xy - k(\beta x^3 + \gamma y^3) = 0.$$

To the planes of a pencil with the line l_1 as axis correspond the points on a cubic of the pencil (38).

[7] E. Čech, *L'intorno di un punto d'una superficie considerato dal punto di vista proiettiva*, Annali di Mathematica, ser. 3, vol. 31(1922), p. 192.

By use of equations (13), (14), we find that the cubic (38) has contact of the second order with each of the projected asymptotic curves at the point P_x in case $k = \frac{1}{2}$. Moreover, inspection of equation (18) shows that the osculating nodal cubic of the projected asymptotics coincides with the cubic (38) if, and only if, the asymptotic curves are projected from a point on the first canonical edge of Green. Thus we arrive at the following theorem:

If the asymptotic curves at a point of a surface are projected upon the tangent plane from a point on the first canonical edge of Green, then to the planes of a pencil with the edge as axis correspond the points on the osculating nodal cubic of the projected asymptotics in the transformation of Čech for which $k = \frac{1}{2}$.

SOUTHWESTERN COLLEGE.

ON ABSOLUTELY CONVERGENT FOURIER-STIELTJES TRANSFORMS

By N. Wiener and H. R. Pitt

(1.1) **Introduction.** Suppose that $f(x)$ is of bounded variation in $(-\infty, \infty)$ and let

$$F(x) = \int e^{-iyx}\,df(y).$$

(Here, and in what follows, integrals in which the limits are not specified are always taken over the range $(-\infty, \infty)$.)

We define $\mathfrak{A}$ to be the class of functions $F(x)$ which can be expressed in this form. If $F(x)$ is periodic, it belongs to $\mathfrak{A}$ only if its Fourier series is absolutely convergent, and Wiener[1] has shown that the condition $F(x) \neq 0$, or the equivalent condition

(1.1.1) $$\underline{\mathrm{Bd}}\,|F(x)| > 0,$$

is sufficient to make $[F(x)]^{-1}$ belong to $\mathfrak{A}$. The same result for almost periodic functions has been proved by Cameron[2] and Pitt.[3]

Our object here is to investigate how far this is true in the general case.

We can write

$$f(x) = h(x) + g(x) + s(x),$$

where $h(x)$ is a step function, $g(x)$ is absolutely continuous and $s(x)$ is singular in the sense defined by Lebesgue; that is, it is continuous, of bounded variation, not constant, and has derivative zero at almost all points. We can write

$$F(x) = H(x) + G(x) + S(x),$$

where $H(x)$, $G(x)$, $S(x)$ are the transforms of $h(x)$, $g(x)$, $s(x)$; and we use the letters H, G, S in this sense throughout the paper. If $F(x)$ belongs to $\mathfrak{A}$ and

$$F(x) = \int e^{-iyx}\,df(y),$$

we write

$$T\{F(x)\} = \int |\,df(y)\,|.$$

Received March 14, 1938.

[1] N. Wiener, *Tauberian theorems*, Annals of Mathematics, vol. 33(1932), pp. 1–100.

[2] R. H. Cameron, *Analytic functions of absolutely convergent generalized trigonometric sums*, this Journal, vol. 3(1937), pp. 682–688.

[3] H. R. Pitt, *A theorem on absolutely convergent trigonometrical series*, Journal of Math. and Phys., M. I. T., vol. 16(1938), pp. 191–195.

Our main conclusion is that the condition $\underline{\mathrm{Bd}} \mid F(x) \mid > 0$ is sufficient to make $[F(x)]^{-1}$ belong to $\mathfrak{A}$ provided that $T\{S(x)\}$ is "not too large", and that otherwise the conclusion is not true.

2. We require a number of elementary lemmas.

(2.1) LEMMA 1. *If $F_1(x)$, $F_2(x)$ belong to $\mathfrak{A}$, then so do $c_1F_1(x) + c_2F_2(x)$ and $F_1(x)F_2(x)$, c_1 and c_2 being any real or complex numbers. Moreover,*

$$T\{c_1F_1(x) + c_2F_2(x)\} \leqq |c_1| \, T\{F_1(x)\} + |c_2| \, T\{F_2(x)\}, \tag{a}$$

$$T\{F_1(x)F_2(x)\} \leqq T\{F_1(x)\} \, T\{F_2(x)\}. \tag{b}$$

The first result is obvious. To prove the second, we write

$$F_1(x) = \int e^{-iyx} \, df_1(y), \qquad F_2(x) = \int e^{-iyx} \, df_2(y).$$

Then

$$\begin{aligned} F_1(x)F_2(x) &= \int e^{-iux} \, df_2(u) \int e^{-iyx} \, df_1(y) \\ &= \int df_2(u) \int e^{-iyx} \, df_1(y - u) \\ &= \int e^{-iyx} \, df(y), \end{aligned}$$

where

$$f(y) = \int f_1(y - u) \, df_2(u),$$

except possibly in an enumerable set of points. If $f(y)$ is defined appropriately at these points,

$$T\{F_1(x)F_2(x)\} = \int |df(y)| \leqq \int |df_1(y)| \int |df_2(y)| = T\{F_1(x)\}T\{F_2(x)\}.$$

(2.2) LEMMA 2. *If $F(x)$ belongs to $\mathfrak{A}$ and $T\{F(x)\} = \alpha < 1$, then $[1 + F(x)]^{-1}$ belongs to $\mathfrak{A}$ and*

$$T\{[1 + F(x)]^{-1}\} \leqq (1 - \alpha)^{-1}.$$

We have

$$[1 + F(x)]^{-1} = \sum_{n=0}^{\infty} (-1)^n [F(x)]^n, \tag{2.2.1}$$

the series being absolutely and uniformly convergent for $-\infty < x < \infty$. It follows from Lemma 1(b) that $[F(x)]^n$ belongs to $\mathfrak{A}$ and that $T\{[F(x)]^n\} \leqq \alpha^n$. We can therefore write

$$[1 + F(x)]^{-1} = \sum_{n=0}^{\infty} \int e^{-iyx} df_n(y), \qquad \int |df_n(y)| \leqq \alpha^n,$$

$$[1 + F(x)]^{-1} = \int e^{-iyx} d\left[\sum_{n=0}^{\infty} f_n(y)\right],$$

$$T\{[1 + F(x)]^{-1}\} = \int \left| d\left[\sum_{n=0}^{\infty} f_n(y)\right]\right| \leqq \sum_{n=0}^{\infty} \int |df_n(y)| \leqq \sum_{n=0}^{\infty} \alpha^n = (1 - \alpha)^{-1}.$$

(2.3) We require certain auxiliary functions $\Omega(\epsilon, \lambda, x)$ and $\Gamma(x)$ which are defined as follows. If $\epsilon > 0$, $\lambda > 0$, $2\epsilon\lambda \leqq \pi$, we write

$$\Omega(\epsilon, \lambda, x) = \begin{cases} 1 & (0 \leqq |x| < \epsilon), \\ 2 - \dfrac{|x|}{\epsilon} & (\epsilon \leqq |x| < 2\epsilon), \\ 0 & (2\epsilon \leqq |x| \leqq \pi/\lambda), \end{cases}$$

$$\Omega(\epsilon, \lambda, x + 2\pi/\lambda) = \Omega(\epsilon, \lambda, x).$$

We write

$$\Gamma(x) = \begin{cases} \Omega(1, \tfrac{1}{2}\pi, x) & (|x| \leqq 2), \\ 0 & (|x| > 2). \end{cases}$$

LEMMA 3. (a) *$\Omega(\epsilon, \lambda, x)$ belongs to $\mathfrak{A}$ and*

$$T\{\Omega(\epsilon, \lambda, x)\} \leqq C,$$

C being an absolute constant.

(b) *For any fixed λ,*

$$\lim_{\epsilon \to 0} T\{\Omega(\epsilon, \lambda, x)[1 - e^{-i\lambda x}]\} = 0.$$

(c) $$\Gamma(x) = \int e^{-iyx}\gamma(y)\, dy, \qquad \int |\gamma(y)|\, dy < \infty.$$

Let us first consider (a). An easy calculation shows that

$$\Omega(\epsilon, \lambda, x) = \sum_{-\infty}^{\infty} d_n e^{-i\lambda nx},$$

where

$$d_0 = \frac{3\lambda\epsilon}{2\pi}, \qquad d_n = \frac{2}{\pi\lambda\epsilon n^2} \sin\frac{n\lambda\epsilon}{2} \sin\frac{3n\lambda\epsilon}{2} \qquad (n \neq 0).$$

It is plain that $\Omega(\epsilon, \lambda, x)$ belongs to $\mathfrak{A}$. Further,

$$T\{\Omega(\epsilon, \lambda, x)\} \leqq \sum_{3\lambda\epsilon|n|\leqq\pi} |d_n| + \frac{2}{\pi\lambda\epsilon} \sum_{3\lambda\epsilon|n|>\pi} \frac{1}{n^2}$$

$$\leqq \frac{3\lambda\epsilon}{2\pi} + 1 + \frac{2}{\pi\lambda\epsilon} \sum_{3\lambda\epsilon|n|>\pi} \frac{1}{n^2} \leqq C,$$

since $2\epsilon\lambda \leqq \pi$.

Next consider (b). We have

$$\Omega(\epsilon, \lambda, x)[1 - e^{-i\lambda x}] = \sum (d_n - d_{n-1})e^{-i\lambda nx}.$$

It is easy to show that $\lim_{\epsilon\to 0} (d_n - d_{n-1}) = 0$ for each n and that

$$|d_n - d_{n-1}| \leqq \frac{A}{n^2} \qquad (n \neq 0,\, n \neq 1),$$

A being independent of ϵ. The conclusion then follows by uniform convergence.

(c) is obvious by direct calculation.

(2.4) Lemma 4. *If $G(x)$ belongs to $\mathfrak{A}$ with $g(x)$ absolutely continuous, and if*

$$\underline{\mathrm{Bd}}\, |1 + G(x)| > 0,$$

then $[1 + G(x)]^{-1}$ belongs to $\mathfrak{A}$.

Let N be a positive integer and let

$$3\epsilon^2 N = 4, \qquad x_n = 3\epsilon n - \frac{2}{\epsilon} \qquad (n = 0, 1, \cdots, N).$$

Then

$$\sum_{n=0}^{N} \Gamma[(x - x_n)/\epsilon] = 1 \qquad (|x| \leqq 2/\epsilon),$$

$$1 = \Gamma(\epsilon x) + 1 - \Gamma(\epsilon x),$$

and since $\Gamma(\epsilon x) = 0$ for $x \geqq 2/\epsilon$, we have

$$1 = \Gamma(\epsilon x) \sum_{n=0}^{N} \Gamma[(x - x_n)/\epsilon] + [1 - \Gamma(\epsilon x)],$$

$$\text{(2.4.1)} \qquad \frac{1}{1 + G(x)} = \Gamma(\epsilon x) \sum_{n=0}^{N} \frac{\Gamma[(x - x_n)/\epsilon]}{1 + G(x)} + \frac{1 - \Gamma(\epsilon x)}{1 + G(x)}.$$

We shall show that each term of this sum belongs to $\mathfrak{A}$ when N is sufficiently large.

We have

$$\frac{\Gamma[(x - x_n)/\epsilon]}{1 + G(x)} = \frac{\Gamma[(x - x_n)/\epsilon]}{1 + G(x_n) + [G(x) - G(x_n)]\Gamma[\frac{1}{2}(x - x_n)/\epsilon]},$$

since the denominators differ only at points where $\Gamma[(x - x_n)/\epsilon] = 0$. Writing $\gamma_\epsilon(y) = 2\epsilon\gamma(2\epsilon y)$, we have

$$[G(x) - G(x_n)]\,\Gamma[\tfrac{1}{2}(x - x_n)/\epsilon]$$

$$= \int e^{-itx} g(t)\,dt \int e^{-iyx+iyx_n}\gamma_\epsilon(y)\,dy - \int e^{-itx_n} g(t)\,dt \int e^{-iyx+iyx_n}\gamma_\epsilon(y)\,dy$$

$$= \int g(t)\,dt \int e^{-iyx+ix_n(y-t)}\gamma_\epsilon(y - t)\,dy - \int e^{-itx_n} g(t)\,dt \int e^{-iyx+iyx_n}\gamma_\epsilon(y)\,dy$$

$$= \int e^{-iyx}\,dy \int e^{-ix_n(y-t)} g(t)[\gamma_\epsilon(y - t) - \gamma_\epsilon(y)]\,dt.$$

Hence

$$T\{[G(x) - G(x_n)]\,\Gamma[\tfrac{1}{2}(x - x_n)/\epsilon]\} \leqq \int \left| \int g(t)[\gamma_\epsilon(y - t) - \gamma_\epsilon(y)]\,dt \right| dy$$

$$\leqq \int | g(t) |\,dt \int | \gamma_\epsilon(y - t) - \gamma_\epsilon(y) |\,dy.$$

But

$$\int | \gamma_\epsilon(y - t) - \gamma_\epsilon(y) |\,dy = \int | \gamma(y - 2\epsilon t) - \gamma(y) |\,dy,$$

and this is bounded for all t and ϵ and tends to zero, for any fixed t, as $\epsilon \to 0$.

Since $| 1 + G(x_n) |$ exceeds a positive number independent of N, it follows that if N is sufficiently large,

$$T\{[G(x) - G(x_n)]\,\Gamma[\tfrac{1}{2}(x - x_n)/\epsilon]\} \leqq | 1 + G(x_n) |$$

for $n = 0, 1, \cdots, N$. Hence, appealing to Lemmas 1 and 2, we see that each term of the first part of the right side of (2.4.1) belongs to $\mathfrak{A}$.

As for the last term on the right of (2.4.1), we can write

$$\frac{1 - \Gamma(\epsilon x)}{1 + G(x)} = \frac{1 - \Gamma(\epsilon x)}{1 + G(x)[1 - \Gamma(\tfrac{1}{2}\epsilon x)]}.$$

Now

$$\Gamma(\tfrac{1}{2}\epsilon x) = \frac{2}{\epsilon} \int e^{-iyx}\gamma(2y/\epsilon)\,dy,$$

$$G(x) = \int e^{-iyx} g(y)\,dy.$$

Therefore

$$G(x)\,\Gamma(\tfrac{1}{2}\epsilon x) = \frac{2}{\epsilon}\int e^{-iyx}\,dy\int g(y-t)\gamma(2t/\epsilon)\,dt,$$

$$\begin{aligned} G(x)[1-\Gamma(\tfrac{1}{2}\epsilon x)] &= \int e^{-iyx}\left[g(y) - \frac{2}{\epsilon}\int g(y-t)\gamma(2t/\epsilon)\,dt\right]dy \\ &= \frac{2}{\epsilon}\int e^{-iyx}\left[\int [g(y)-g(y-t)]\gamma(2t/\epsilon)\,dt\right]dy, \end{aligned}$$

since

$$\frac{2}{\epsilon}\int \gamma(2t/\epsilon)\,dt = 1.$$

Hence

$$\begin{aligned} T\{G(x)[1-\Gamma(\tfrac{1}{2}\epsilon x)]\} &= \frac{2}{\epsilon}\int\int |\,[g(y)-g(y-t)]\gamma(2t/\epsilon)\,|\,dt\,dy \\ &= \int |\,\gamma(t)\,|\,dt\int |\,g(y)-g(y-\tfrac{1}{2}t\epsilon)\,|\,dy. \end{aligned}$$

Since $g(y) \in L(-\infty, \infty)$, $\int |\,g(y) - g(y-\frac{1}{2}t\epsilon)\,|\,dy$ is bounded for $-\infty < t < \infty$ and tends to zero for each fixed t, as $\epsilon \to 0$. Hence

$$\lim_{\epsilon\to 0} T\{G(x)[1-\Gamma(\tfrac{1}{2}\epsilon x)]\} = 0,$$

and it follows from Lemmas 1 and 2 that $[1 - \Gamma(\epsilon x)]/[1 + G(x)]$ belongs to $\mathfrak{A}$ when ϵ is sufficiently small.

(2.5) Lemma 5. *Let $G(x)$, $F(x)$, $[F(x)]^{-1}$ belong to $\mathfrak{A}$, $g(x)$ being absolutely continuous. Let F and G satisfy*

$$\underline{\mathrm{Bd}}\,|\,F(x)+G(x)\,| > 0.$$

Then $[F(x)+G(x)]^{-1}$ belongs to $\mathfrak{A}$.

We have

$$\frac{1}{F(x)+G(x)} = \frac{1}{F(x)}\,\frac{1}{\{1+G(x)[F(x)]^{-1}\}}.$$

Let

$$[F(x)]^{-1} = \int e^{-iyx}\,df^*(y), \qquad \int |\,df^*(y)\,| < \infty.$$

Since $g(x)$ is absolutely continuous, we can write

$$G(x) = \int e^{-iyx}g'(y)\,dy, \qquad \int |\,g'(y)\,|\,dy < \infty.$$

Then

$$G(x)[F(x)]^{-1} = \int e^{-iyx}\,dy \int g'(y - t)\,df^*(t),$$

and since $F(x)$ is bounded,

$$\underline{\mathrm{Bd}}\,|\,1 + G(x)[F(x)]^{-1}\,| = \underline{\mathrm{Bd}}\,|\,[F(x) + G(x)][F(x)]^{-1}\,| > 0,$$

and the conclusion follows from Lemma 4.

(3.1) THEOREM 1. *Suppose that* $F(x)$ *belongs to* $\mathfrak{A}$ *and let* $\underline{\mathrm{Bd}}\,|\,F(x)\,| > 0$. *Suppose that* $H(x)$, $S(x)$ *are defined as in* (1.1), *and that*

$$T\{S(x)\} < \underline{\mathrm{Bd}}\,|\,H(x)\,|.$$

Then $[F(x)]^{-1}$ *belongs to* $\mathfrak{A}$.

In view of Lemma 5, it is sufficient to prove that $[H(x) + S(x)]^{-1}$ belongs to $\mathfrak{A}$. Let

$$\begin{aligned} &\underline{\mathrm{Bd}}\,|\,H(x)\,| = \delta, \\ &0 < 3\eta \leqq \delta - T\{S(x)\}. \end{aligned} \tag{3.1.1}$$

Then

$$H(x) = H_1(x) + H_2(x),$$

where

$$H_1(x) = \sum_{n=1}^{N} a_n e^{-i\lambda_n x}, \qquad H_2(x) = \sum_{n=N+1}^{\infty} a_n e^{-i\lambda_n x},$$

and N is chosen so that

$$T\{H_2(x)\} = \sum_{n=N+1}^{\infty} |\,a_n\,| \leqq \eta, \tag{3.1.2}$$

$$|\,H_1(x)\,| \geqq \delta - \eta. \tag{3.1.3}$$

Let K be an integer greater than 1 and let

$$\epsilon_n = \frac{2\pi}{3K\lambda_n}, \qquad \Omega_n(x) = \Omega(\epsilon_n, \lambda_n, x)$$

for $n = 1, 2, \cdots, N$. Then

$$\sum_{k=1}^{K} \Omega_n(x - 3k\epsilon_n) = 1$$

for $n = 1, 2, \cdots, N$ and all x, and it follows that

$$\prod_{n=1}^{N} \sum_{k=1}^{K} \Omega_n(x - 3k\epsilon_n) = 1.$$

Hence

$$\sum \prod_{n=1}^{N} \Omega_n(x - 3k_n \epsilon_n) = 1,$$

where the summation extends over the K^N sets $(k_1, \cdots, k_N)$ formed from integers $1, 2, \cdots, K$.

We can therefore write

$$\frac{1}{H(x) + S(x)} = \sum \frac{\prod_{n=1}^{N} \Omega_n(x - 3k_n \epsilon_n)}{H(x) + S(x)}. \tag{3.1.4}$$

By Lemma 1, it is sufficient to show that each term of the sum belongs to $\mathfrak{A}$. Let x_0 be a point at which

$$\Omega(x) = \prod_{n=1}^{N} \Omega_n(x - 3k_n \epsilon_n) \neq 0.$$

(If there is no such point, the relevant term in the right side of (3.1.4) vanishes.) Then

$$\frac{\Omega(x)}{H(x) + S(x)} = \frac{\Omega(x)}{H_1(x_0) + Q(x)}, \tag{3.1.5}$$

where

$$Q(x) = H_2(x) + S(x) + [H_1(x) - H_1(x_0)] \prod_{n=1}^{N} \Omega(2\epsilon_n, \lambda_n, x - 3\epsilon_n k_n), \tag{3.1.6}$$

since the two denominators differ only at points where $\Omega(x) = 0$. Since

$$[H_1(x) - H_1(x_0)] \prod_{n=1}^{N} \Omega(2\epsilon_n, \lambda_n, x - 3\epsilon_n k_n)$$
$$= \sum_{\nu=1}^{N} a_\nu (e^{-i\lambda_\nu x} - e^{-i\lambda_\nu x_0}) \prod_{n=1}^{N} \Omega(2\epsilon_n, \lambda_n, x - 3\epsilon_n k_n),$$

we have

$$\begin{aligned} &T\left\{[H_1(x) - H_1(x_0)] \prod_{n=1}^{N} \Omega(2\epsilon_n, \lambda_n, x - 3\epsilon_n k_n)\right\} \\ &\leqq \sum_{\nu=1}^{N} |a_\nu| C^{N-1} T\{[e^{-i\lambda_\nu x} - e^{-i\lambda_\nu x_0}]\Omega(2\epsilon_\nu, \lambda_\nu, x - 3\epsilon_\nu k_\nu)\}, \end{aligned} \tag{3.1.7}$$

by Lemmas 1 and 3. Since $\Omega(x_0) \neq 0$, we have

$$\Omega_\nu(x_0 - 3\epsilon_\nu k_\nu) \neq 0 \qquad (\nu = 1, 2, \cdots, N),$$

and therefore, by the definition of $\Omega_\nu(x)$,

$$|x_0 - 3\epsilon_\nu k_\nu| < 2\epsilon_\nu, \qquad \mathrm{mod}\ (2\pi/\lambda_\nu), \tag{3.1.8}$$

for $\nu = 1, 2, \cdots, N$. Returning to (3.1.7), we see that

$$
\begin{aligned}
&T\{[e^{-i\lambda_\nu x} - e^{-i\lambda_\nu x_0}]\,\Omega(2\epsilon_\nu, \lambda_\nu, x - 3\epsilon_\nu k_\nu)\} \\
(3.1.9)\qquad &= T\{[e^{-i\lambda_\nu(x+3\epsilon_\nu k_\nu)} - e^{-i\lambda_\nu x_0}]\,\Omega(2\epsilon_\nu, \lambda_\nu, x)\} \\
&= T\{[e^{-i\lambda_\nu x} - e^{-i\lambda_\nu(x_0-3\epsilon_\nu k_\nu)}]\,\Omega(2\epsilon_\nu, \lambda_\nu, x)\} \\
&\leqq T\{[1 - e^{-i\lambda_\nu x}]\Omega(2\epsilon_\nu, \lambda_\nu, x)\} + C\,|\,1 - e^{-i\lambda_\nu(x_0-3\epsilon_\nu k_\nu)}\,|.
\end{aligned}
$$

Hence, using (3.1.7), (3.1.8) and Lemma 1, we can choose K so that

$$(3.1.10)\qquad T\left\{[H_1(x) - H_1(x_0)]\prod_{n=1}^{N}\Omega(2\epsilon_n, \lambda_n, x - 3\epsilon_n k_n)\right\} < \eta.$$

Then, by (3.1.1), (3.1.2), (3.1.3) and (3.1.10), we have

$$T\{Q(x)\} < \eta + \delta - 3\eta + \eta = \delta - \eta \leqq H_1(x_0),$$

and the conclusion follows from (3.1.5) by using Lemma 2.

(4.1) We shall devote the remainder of this paper to showing that Theorem 1 is substantially the best possible result of its kind. In particular, we show that the conditions

$$\underline{\mathrm{Bd}}\,|\,F(x)\,| > 0, \qquad F(x) \,\epsilon\, \mathfrak{A}$$

are not sufficient to make $[F(x)]^{-1} \,\epsilon\, \mathfrak{A}$. We require Theorem 2 in the proof of Theorem 3, but it seems of some independent interest in itself.

THEOREM 2. *Let δ be > 0. Then we can define a function $S(x)$ of $\mathfrak{A}$ such that*

$$(4.2.1)\qquad |\,S(x)\,| \leqq \delta,$$

$$(4.2.2)\qquad T\{S(x)\} = 1,$$

$$(4.2.3)\qquad T\left\{\sum_{k=1}^{\infty} H_k(x)[S(x)]^k\right\} = \sum_{k=1}^{\infty} T\{H_k(x)\}$$

for any sequence of functions $H_k(x)$ defined by

$$(4.2.4)\qquad H_k(x) = \sum_{j=1}^{\infty} a_{k,j}e^{-i\lambda_j x}, \qquad \sum_{k,j}|\,a_{k,j}\,| < \infty.$$

We observe, first, that if $S(x)$ satisfies conditions (4.2.2), (4.2.3), then so does $[S(x)]^m$ for any integer m. Therefore, instead of (4.2.1), it is sufficient to show that

$$(4.2.5)\qquad \overline{\mathrm{Bd}}\,|\,S(x)\,| < 1,$$

for then $[S(x)]^m$ will be a function of the required type if m is sufficiently large.

In what follows we suppose that the function $s(x)$ is of bounded variation in $(-\infty, \infty)$ and define $S(x)$ by

$$S(x) = \int e^{-iyx}\,ds(y).$$

We denote by B (with or without a suffix) any array of the form

$$\cdot\ ---\ 00\ ---\ xxx\ ---\ 00-\ \cdots$$

in which there are arbitrarily long stretches of consecutive 0's, x's, and gaps. We shall define $s(x)$ by means of such arrays.

With any zero of B, at the r-th place say, we associate the set of 2^{r-1} intervals of length 2^{-r} whose binary representations have zeros in the r-th place, and suppose that $s(x)$ is constant in the complementary intervals. We suppose also that whenever x occurs in the r-th place of B, the variation (possibly zero) of $s(x)$ in each interval $(\nu 2^{-r}, (\nu + 1)2^{-r})$ $(\nu = 0, 1, \cdots, 2^r - 1)$ is divided equally between the two intervals $(\nu 2^{-r}, \nu 2^{-r} + 2^{-r-1})$ and $(\nu 2^{-r} + 2^{-r-1}, (\nu + 1) 2^{-r})$. In these circumstances we say that the function $e^{-i\lambda x}S(x)$, for any real λ, is of type B.

We say that two arrays are *disjoint* if there are arbitrarily long stretches of consecutive 0's in each which correspond to x's in the other. A set of arrays is disjoint if each pair is disjoint. We require the following result.

LEMMA 6. *Let $B_1, \cdots, B_K$ be disjoint, and let $S_1(x), \cdots, S_K(x)$ be of types $B_1, \cdots, B_K$, respectively. Let $H_k(x)$ $(k = 1, \cdots, K)$ be trigonometrical polynomials. Then*

$$T\left\{\sum_{k=1}^{K} S_k(x)H_k(x)\right\} = \sum_{k=1}^{K} T\{S_k(x)H_k(x)\}.$$

We may suppose without loss of generality that each of the polynomials $H_k(x)$ has K terms. Suppose that $S_\nu(x)$ is of type B_ν $(\nu = 1, \cdots, K)$, and suppose that B_1 has 0's and B_2 has x's from the N-th to the $(N + p)$-th place. We write

$$S_\nu(x)H_\nu(x) = \int e^{-iyx}\, d\sigma_\nu(y) \qquad (\nu = 1, \cdots, K).$$

Then the whole variation of $\sigma_1(y)$ is concentrated in K sets of 2^N intervals of the form

$$(\lambda_k + \nu 2^{-N}, \lambda_k + \nu 2^{-N} + 2^{-N-p}) \qquad (k = 1, \cdots, K; \nu = 0, \cdots, 2^N),$$

where λ_k is an exponent in $H_1(x)$. But for each k, this set of intervals cannot contain more than a fraction 2^{-p+1} of the variation of $s_2(x - \lambda)$, for any λ, since B_2 has x's between the N-th and $(N + p)$-th places. Therefore not more than a fraction $K \cdot 2^{-p+1}$ of the variation of $\sigma_2(x)$ is contained in these intervals. Then plainly

$$T\{S_1(x)H_1(x) + S_2(x)H_2(x)\} \geqq T\{S_1(x)H_1(x)\} + (1 - K2^{-p+1})T\{S_2(x)H_2(x)\},$$

and since p is arbitrarily large,

$$T\{S_1(x)H_1(x) + S_2(x)H_2(x)\} = T\{S_1(x)H_1(x)\} + T\{S_2(x)H_2(x)\}.$$

The extension of this argument to K functions is obvious.

We can now define $S(x)$. We suppose that $s(x)$ increases from 0 to $\frac{1}{2}$ in $(0, \frac{1}{2})$ and decreases from $\frac{1}{2}$ to 0 in $(\frac{1}{2}, 1)$. Let

$$S(x) = \int_0^1 e^{-iyx}\, ds(y).$$

If N is a positive integer, we can write

$$S(x) = \sum_{\nu=0}^{2^N-1} S_\nu^N(x), \qquad S_\nu^N(x) = \int_{\nu 2^{-N}}^{(\nu+1)2^{-N}} e^{-iyx}\, ds(y). \tag{4.2.6}$$

If

$$\alpha_\nu^N = \int_{\nu 2^{-N}}^{(\nu+1)2^{-N}} |\, ds(y)\,|,$$

then

$$\sum_{\nu=0}^{2^N-1} \alpha_\nu^N = 1.$$

We proceed inductively (with respect to N). Suppose that $S_\nu^N(x)$ is of type B_ν^N, that the B_ν^N's have arbitrarily long stretches of gaps in common and that they have 0's or x's in the first N places. Moreover, we suppose that the B_ν^N's are such that if they are divided into two groups in any way, the arrays in each group have in common arbitrarily long stretches of 0's corresponding to x's in all the arrays of the other group.

Now consider the two groups of integers ν for which B_ν^N has x's or 0's, respectively, in the n-th place (these may not exhaust the integers $0, 1, \cdots, 2^N - 1$, as some of the B_ν^N's may have gaps). We suppose that for every n, the sum of α_ν^N's of the first group is not less than the sum of α_ν^N's of the second. In other words, we suppose that at least half of the variation of $s(x)$ is distributed by an x at each stage.

Now since we have only a finite number of B_ν^N's to define for each value of N, and we have arbitrarily long stretches of gaps in the arrays defined for integers less than N, it is plain that the B_ν^N's can be defined inductively so that the properties described above hold for every N. Since this process defines α_ν^N for all N and ν, the function $s(x)$ is defined completely.

Since $T\{[S(x)]^k\} \leqq 1$, it is sufficient to prove the theorem in the case in which there is a finite number of $H_k(x)$'s and each is a trigonometric polynomial. The result in the general case follows at once by a simple limit argument. Let ϵ be > 0, and let 2^N be $> k$. Then

$$\begin{aligned} &S(x) = \sum_\nu S_\nu^N(x), \\ &[S(x)]^k = [\sum_\nu S_\nu^N(x)]^k = k! \sum S_{\nu_1}^N(x) S_{\nu_2}^N(x) \cdots S_{\nu_k}^N(x) + Q_k^N(x), \end{aligned} \tag{4.2.7}$$

where $\nu_1, \nu_2, \cdots, \nu_k$ are all different and $Q_k^N(x)$ is the sum of terms involving the square or higher power of at least one $S_\nu^N(x)$.

It is plain from the construction of $S(x)$ that

$$\lim_{N\to\infty} \max \alpha_\nu^N = 0,$$

and since $s(y)$ is monotonic in each interval $(\nu 2^{-N}, (\nu + 1)2^{-N})$,

$$T\{S_{\nu_1}^N(x) \cdots S_{\nu_k}^N(x)\} = T\{S_{\nu_1}^N(x)\} \cdots T\{S_{\nu_k}^N(x)\} = \alpha_{\nu_1}^N \cdots \alpha_{\nu_k}^N.$$

Hence[4]

$$\text{(4.2.8)} \qquad \lim_{N\to\infty} T\{Q_k^N(x)\} = 0, \qquad \lim_{N\to\infty} k! \sum T\{S_{\nu_1}^N(x) \cdots S_{\nu_k}^N(x)\} = 1.$$

Now if two functions $S(x)$, $S'(x)$ are of types B, B', it is easy to show that $S(x)S'(x)$ is of a type obtained by adding B and B' according to the rule that in stretches of consecutive 0's or x's we write

$$0 + 0 = 0, \qquad x + 0 = x, \qquad x + x = x,$$

except possibly in the last place in any stretch. Thus, if

$$B = \cdot 0000 --- xxxx --- xxxx ---,$$

$$B' = \cdot 00000 -- 000 ---- xxxx ---,$$

then $S(x)S'(x)$ is of type

$$\cdot 000 ---- xx ------ xxx ----.$$

Then it follows from our definition of $S_\nu^N(x)$ that non-identical functions of the form

$$S_{\nu_1}^N(x) S_{\nu_2}^N(x) \cdots S_{\nu_k}^N(x)$$

are disjoint, so that, using (4.2.7) and (4.2.8), we can write

$$\text{(4.2.9)} \qquad [S(x)]^k = \sum_p S_{k,p}(x) + Q_k(x),$$

where all the $S_{k,p}(x)$'s are disjoint and

$$\text{(4.2.10)} \qquad \sum_{k,j} |a_{k,j}| \, T\{Q_k(x)\} \leqq \epsilon,$$

$$\text{(4.2.11)} \qquad \sum_p T\{S_{k,p}(x)\} \geqq 1 - \epsilon \qquad (k = 1, 2, \cdots, K).$$

[4] We use the fact that if $\alpha_\nu \geqq 0$, $\sum_{\nu=1}^{n} \alpha_\nu = 1$, then

$$(\alpha_1 + \alpha_2 + \cdots + \alpha_n)^k - k! \sum \alpha_{\nu_1}\alpha_{\nu_2} \cdots \alpha_{\nu_k} \to 0$$

as $\max \alpha_\nu \to 0$, the summation being over all sets of k different ν_i's.

Hence

$$\sum_k H_k(x)[S(x)]^k = \sum_k \sum_j a_{k,j} e^{-i\lambda_j x} [\sum_p S_{k,p}(x) + Q_k(x)],$$

$$T\{\sum_k H_k(x)[S(x)]^k\} \geqq T\{\sum_k \sum_p \sum_j a_{k,j} e^{-i\lambda_j x} S_{k,p}(x)\} - \epsilon$$

$$= \sum_{k,p} T\{\sum_j a_{k,j} e^{-i\lambda_j x} S_{k,p}(x)\} - \epsilon,$$

by Lemma 6. Now

$$S_{k,p}(x) = \int e^{-iyx}\, ds_{k,p}(y),$$

where $s_{k,p}(x)$ is constant except in an interval of length $k \cdot 2^{-N}$. Hence, if we choose N so that

$$k \cdot 2^{-N} \leqq \min_{j \neq j'} |\lambda_j - \lambda_{j'}|,$$

it is plain that

$$T\{\sum_j a_{k,j} e^{-i\lambda_j x} S_{k,p}(x)\} = \sum_j |a_{k,j}|\, T\{e^{-i\lambda_j x} S_{k,p}(x)\}$$

$$= \sum_j |a_{k,j}|\, T\{S_{k,p}(x)\}.$$

It follows that

$$T\{\sum_k H_k(x)[S(x)]^k\} \geqq \sum_{k,j} |a_{k,j}| \sum_p T\{S_{k,p}(x)\} - \epsilon$$

$$\geqq \sum_{k,j} |a_{k,j}| - \epsilon[1 + \sum_{k,j} |a_{k,j}|],$$

by (4.2.11). Since ϵ is arbitrary,

$$T\{\sum_k H_k(x)[S(x)]^k\} \geqq \sum_{k,j} |a_{k,j}| = \sum_k T\{H_k(x)\},$$

and since

$$T\{\sum_k H_k(x)[S(x)]^k\} \leqq \sum_k T\{H_k(x)\},$$

by Lemma 1, this completes the proof of (4.2.3).

Finally, we have to prove (4.2.5). Let x be $\geqq \frac{1}{2}\pi$, and choose an integer $N \geqq 1$ so that

$$\pi \cdot 2^{N-2} \leqq x < \pi \cdot 2^{N-1}. \tag{4.2.12}$$

We write, as before,

$$S(x) = \sum_{\nu=0}^{2^N-1} S_\nu^N(x).$$

Since at least a half of the total variation of $s(x)$ is divided equally between intervals $(\nu 2^{-N}, \nu 2^{-N} + 2^{-N-1})$ and $(\nu 2^{-N} + 2^{-N-1}, (\nu + 1)2^{-N})$, we can write

$$|S_\nu^N(x)| \leqq \beta_\nu + \sum_\nu \left| \int_0^{2^{-N}} e^{-iyx}\, d\sigma_\nu(y) \right|,$$

where $\beta_\nu \geqq 0$, $\sigma_\nu(y)$ increases from 0 to β'_ν in $(0, 2^{-N})$,

$$0 \leqq A \leqq \tfrac{1}{2}, \qquad \sum_\nu \beta_\nu = A, \qquad \sum_\nu \beta'_\nu = 1 - A.$$

Moreover, since the variation of $s(y)$ in the interval $(\nu 2^{-M}, \nu 2^{-M} + 2^{-M-k})$ is at least a fraction 2^{-k} of its variation in $(\nu 2^{-M}, (\nu + 1)2^{-M})$, we can suppose that

$$\sigma_\nu(2^{-N-3}) = \lambda_1 \beta'_\nu \geqq \tfrac{1}{8}\beta'_\nu,$$

$$\sigma_\nu(2^{-N-1} + 2^{-N-3}) - \sigma_\nu(2^{-N-1}) = \lambda_2 \beta'_\nu \geqq \tfrac{1}{8}\beta'_\nu.$$

Hence

$$\left| \int_0^{2^{-N}} e^{-iyx}\, d\sigma_\nu(y) \right| \leqq (1 - \lambda_1 - \lambda_2)\beta'_\nu + \left| \int_0^{2^{-N-3}} e^{-iyx}\, d\sigma_\nu(y) + \int_{2^{-N-1}}^{2^{-N-1}+2^{-N-3}} e^{-iyx}\, d\sigma_\nu(y) \right|.$$

Using (4.2.12), we see that

$$0 \leqq xy \leqq \tfrac{1}{16}\pi$$

in the first integral, while

$$\tfrac{1}{8}\pi \leqq xy \leqq \tfrac{5}{16}\pi$$

in the second. It follows that

$$\left| \int_0^{2^{-N}} e^{-iyx}\, d\sigma_\nu(y) \right| \leqq (1 - \lambda_1 - \lambda_2)\beta'_\nu + \beta'_\nu[\lambda_1 + \lambda_2 - \tfrac{1}{4} + \tfrac{1}{8}\,|1 + e^{i\theta_\nu}|]$$
$$= \beta'_\nu[\tfrac{3}{4} + \tfrac{1}{8}\,|1 + e^{i\theta_\nu}|],$$

where

$$\tfrac{1}{16}\pi \leqq \theta_\nu \leqq \tfrac{5}{16}\pi.$$

Hence

$$\left| \int_0^{2^{-N}} e^{-iyx}\, d\sigma_\nu(y) \right| \leqq C\beta'_\nu,$$

C being an absolute constant less than 1, and we have

$$|S(x)| \leqq A + (1 - A)C < 1$$

for $x \geqq \frac{1}{2}\pi$. If $x < \frac{1}{2}\pi$, we use the fact that $s(y)$ increases from 0 to $\frac{1}{2}$ in $(0, \frac{1}{2})$ and decreases from $\frac{1}{2}$ to 0 in $(\frac{1}{2}, 1)$. We have

$$S(x) = Ae^{i\theta} - Be^{i\varphi},$$

where

$$0 \leqq A \leqq \tfrac{1}{2}, \qquad 0 \leqq B \leqq \tfrac{1}{2},$$

$$0 \leqq \theta(x) \leqq \tfrac{1}{2}x, \qquad \tfrac{1}{2}x \leqq \varphi(x) \leqq x.$$

Then we have

$$0 \leqq \psi(x) = \varphi(x) - \theta(x) \leqq \tfrac{1}{2}\pi,$$

$$\begin{aligned} | S(x) | &\leqq | A - Be^{i\psi} | \\ &= | A - B \cos \psi - iB \sin \psi | \\ &\leqq | A - iB | \leqq 2^{-\frac{1}{2}} \end{aligned}$$

This completes the proof of Theorem 2.

(5.1) THEOREM 3. *Let δ, δ' satisfy $\delta > \delta' > 0$. Then there is a function $S(x)$ of $\mathfrak{A}$, with*

$$| S(x) | \leqq \delta', \qquad T\{S(x)\} = \delta,$$

such that, if $\underline{\mathrm{Bd}}\, | H(x) | = \delta$,

$$\frac{1}{F(x)} = \frac{1}{H(x) + S(x) + G(x)}$$

does not belong to $\mathfrak{A}$ for any $G(x)$ (even if $\underline{\mathrm{Bd}}\, | F(x) | > 0$).

If $| S(x) | \leqq \delta'$, then $| H(x) + S(x) | \geqq \delta - \delta' > 0$, so that it is sufficient, after Theorem 1, to show that $[H(x) + S(x)]^{-1}$ does not belong to $\mathfrak{A}$. We may plainly suppose that $\delta = 1$ and that

$$\underline{\mathrm{Bd}}\, | H(x) - 1 | = 0. \tag{5.1.1}$$

Let η be > 0, and let

$$H(x) = \sum_{n=1}^{\infty} a_n e^{-i\lambda_n x},$$

$$H_1(x) = \sum_{n=1}^{N} a_n e^{-i\lambda_n x}, \qquad H_2(x) = \sum_{N+1}^{\infty} a_n e^{-i\lambda_n x},$$

where $N = N(\eta)$ is chosen so that

$$T\{H_2(x)\} \leqq \tfrac{1}{4}\eta, \qquad | \underline{\mathrm{Bd}}\, | H_1(x) | - 1 | \leqq \tfrac{1}{4}\eta. \tag{5.1.2}$$

Now let $S(x)$ be the function defined in Theorem 2 (with δ' instead of δ). We assume that $[H(x) + S(x)]^{-1}$ belongs to $\mathfrak{A}$, and show that this leads to a contradiction. Let

$$T\{[H(x) + S(x)]^{-1}\} = A < \infty. \tag{5.1.3}$$

Using (5.1.1) and the first inequality of (5.1.2), we can choose x_0 so that

$$| H_1(x_0) - 1 | \leqq \tfrac{1}{2}\eta. \tag{5.1.4}$$

By Lemma 3(b), we can choose $\epsilon = \epsilon(N, \eta) = \epsilon(\eta)$ so that

$$T\{[H_1(x) - H_1(x_0)] \prod_{n=1}^{N} \Omega(2\epsilon, \lambda_n, x - x_0)\} < \tfrac{1}{4}\eta.$$

Then since

$$H(x) + S(x) = H_1(x_0) + H_2(x) + S(x) + [H_1(x) - H_1(x_0)] \prod_{n=1}^{N} \Omega(2\epsilon, \lambda_n, x - x_0)$$

at points where

$$\Omega(x) = \prod_{n=1}^{N} \Omega(\epsilon, \lambda_n, x - x_0) \neq 0,$$

we have, using (5.1.2), (5.1.4) and Lemma 1,

$$\frac{\Omega(x)}{H(x) + S(x)} = \frac{\Omega(x)}{1 + S(x) + P(x)}, \tag{5.1.5}$$

where

$$T\{P(x)\} \leqq \eta. \tag{5.1.6}$$

Now let M be a positive integer independent of η and let

$$M > 2A - 1. \tag{5.1.7}$$

We have

$$\begin{aligned} \frac{1}{1 + S(x) + P(x)} &= \sum_{\nu=0}^{M} (-1)^\nu [S(x) + P(x)]^\nu \\ &\quad + \frac{(-1)^{M+1}[S(x) + P(x)]^{M+1}}{1 + S(x) + P(x)} \\ &= \sum_{\nu=0}^{M} (-1)^\nu [S(x)]^\nu + Q(x), \end{aligned} \tag{5.1.8}$$

where

$$Q(x) = \frac{(-1)^{M+1}[S(x) + P(x)]^{M+1}}{1 + S(x) + P(x)} + \sum_{\nu=0}^{M} (-1)^\nu \{[S(x) + P(x)]^\nu - [S(x)]^\nu\}.$$

Using (5.1.6) and Lemma 1, we obtain

$$T\{S(x) + P(x)\} \leqq 1 + \eta,$$
$$T\{[S(x) + P(x)]^\nu - [S(x)]^\nu\} \leqq (1 + \eta)^\nu - 1.$$

Hence

$$\begin{aligned} T\{Q(x)\Omega(x)\} &\leqq (1 + \eta)^{M+1} T\{\Omega(x)[1 + S(x) + P(x)]^{-1}\} \\ &\quad + T\{\Omega(x)\} \sum_{\nu=0}^{M} [(1 + \eta)^\nu - 1] \\ &\leqq T\{\Omega(x)\} \left[(1 + \eta)^{M+1} A + \sum_{\nu=0}^{M} [(1 + \eta)^\nu - 1] \right], \end{aligned}$$

by (5.1.5) and (5.1.3), and it follows from this and (5.1.5) and (5.1.8) that

$$\begin{aligned} T\left\{\Omega(x)\sum_{\nu=0}^{M}(-1)^{\nu}[S(x)]^{\nu}\right\} &\leqq T\{\Omega(x)[H(x)+S(x)]^{-1}\}+T\{\Omega(x)Q(x)\} \\ &\leqq T\{\Omega(x)\}\left[A(1+(1+\eta)^{M+1})+\sum_{\nu=0}^{M}[(1+\eta)^{\nu}-1]\right]. \end{aligned} \tag{5.1.9}$$

But by Theorem 2, we have

$$T\left\{\Omega(x)\sum_{\nu=0}^{M}(-1)^{\nu}[S(x)]^{\nu}\right\}=\sum_{\nu=0}^{M}T\{\Omega(x)(-1)^{\nu}\}=(M+1)T\{\Omega(x)\}.$$

Combining this with (5.1.9), and dividing by $T\{\Omega(x)\}$, we have

$$M+1\leqq A[1+(1+\eta)^{M+1}]+\sum_{\nu=0}^{M}[(1+\eta)^{\nu}-1].$$

Since M is independent of η we can let $\eta\to 0$ and obtain

$$M+1\leqq 2A.$$

This contradicts (5.1.7). Therefore $[H(x)+S(x)]^{-1}$ cannot belong to $\mathfrak{A}$.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND HARVARD UNIVERSITY.

A REMARK ON WIENER'S GENERAL TAUBERIAN THEOREM

By H. R. Pitt

1. The following theorem[1] is due to Wiener.

Theorem A. *Hypothesis:*[2]

(a) $k(x)$, $k^*(x)$ *belong to* $L(-\infty, \infty)$,

(b) $\int k(x)dx = \int k^*(x)dx = 1$,

(c) $|s(x)| \leqq C$,

(d) $K(x) = \int e^{-iyx}k(y)\,dy \neq 0 \qquad (-\infty < x < \infty)$,

(e) $\lim_{x\to\infty} \int k(x-y)s(y)\,dy = A$.

Conclusion:

$$\lim_{x\to\infty} \int k^*(x-y)s(y)\,dy = A.$$

Our object here is to determine extra conditions on $k(x)$ and $k^*(x)$ under which condition (c) may be replaced by the one-sided condition

$$s(x) \geqq -C.$$

(To show that Theorem A fails with this weaker condition, we take $s(x) = e^{\sigma x}$ $(\sigma > 0)$ and suppose that $k(x)$, $k(x)e^{-\sigma x}$ belong to L and $\int k(x)e^{-\sigma x}dx = 0$.) The result may be stated as follows.

Theorem. *Hypothesis:*

(a) $k(x) \geqq 0$, $k(x)$ *belongs to* $L(-\infty, \infty)$, $\int k(x)dx = 1$,

(b) $k^*(x)$ *is continuous almost everywhere and*

$$\sum_{n=-\infty}^{\infty} \overline{\text{bd}}_{n \leqq x < n+1} |k^*(x)| < \infty, \qquad \int k^*(x)\,dx = 1,$$

(c) $s(x) \geqq -C$,

(d) $K(x) = \int e^{-iyx}k(y)dy \neq 0 \qquad (-\infty < x < \infty)$,

Received May 12, 1938.

[1] N. Wiener, *Tauberian theorems*, Annals of Mathematics, vol. 33(1932), pp. 1–100. The theorem stated here is Wiener's Theorem VIII.

[2] Integrals in which limits are not specified are over the range $(-\infty, \infty)$.

(e) $\int k(x - y)s(y)dy = g(x)$ *exists for every real* x, *is bounded, and*

$$\lim_{x\to\infty} g(x) = A.$$

Conclusion:

$$\lim_{x\to\infty} \int k^*(x - y)s(y)\,dy = A.$$

We may plainly suppose that $A = 0$. First, we show that it is sufficient to prove the theorem when $k(x)$ is continuous. We write

$$h(x) = \pi^{-\frac{1}{2}} \int e^{-y^2} k(x - y)dy,$$

and it is clear that $h(x)$ is continuous and

$$h(x) \geqq 0, \qquad h(x) \in L(-\infty, \infty), \qquad \int h(x)\,dx = 1.$$

Also, since $h(x)$, $k(x)$ are non-negative, and $s(x)$ is bounded below,

$$\begin{aligned}\int h(x - y)s(y)\,dy &= \pi^{-\frac{1}{2}} \int s(y)\,dy \int e^{-t^2} k(x - y - t)\,dt\\ &= \pi^{-\frac{1}{2}} \int e^{-t^2}\,dt \int k(x - t - y)s(y)\,dy\\ &= \pi^{-\frac{1}{2}} \int e^{-t^2} g(x - t)\,dt,\end{aligned}$$

and therefore

$$\lim_{x\to\infty} \int h(x - y)s(y)\,dy = 0.$$

Moreover,

$$\begin{aligned}\int e^{-iyx} h(y)\,dy &= \pi^{-\frac{1}{2}} \int e^{-iyx} k(y)\,dy \int e^{-iyx} e^{-y^2}\,dy\\ &= K(x)e^{-\frac{1}{4}x^2}\\ &\neq 0 \qquad\qquad (-\infty < x < \infty).\end{aligned}$$

The conditions of the theorem are therefore satisfied if $k(x)$ is replaced by the continuous function $h(x)$, so that there is no loss of generality in supposing that $k(x)$ itself is continuous.

Now let δ be > 0 and let

$$S_\delta(x) = \frac{1}{\delta} \int_x^{x+\delta} s(y)\,dy = \frac{1}{\delta} \int_0^\delta s(x + y)\,dy.$$

Let $k(x_0)$ be $> \epsilon > 0$, and define $\delta = \delta(\epsilon)$ so that

$$k(x) \geqq \tfrac{1}{2}\epsilon \qquad (x_0 - \delta \leqq x \leqq x_0).$$

Then since $s(x) + C \geqq 0$, $k(y) \geqq 0$,

$$\begin{aligned} g(x_0 + x) + C &= \int k(y)[s(x + x_0 - y) + C]\,dy \\ &\geqq \tfrac{1}{2}\epsilon \int_{x_0-\delta}^{x_0} [s(x + x_0 - y) + C]\,dy \\ &\geqq \tfrac{1}{2}\epsilon\delta S_\delta(x), \end{aligned}$$

and it follows that $S_\delta(x)$ is bounded for any positive δ. If $D > \delta > 0$, we have

$$\begin{aligned} \int_0^{D-\delta} [S_\delta(x + y) + C]\,dy &\leqq \int_0^D [s(x + y) + C]\,dy \\ &\leqq \int_0^{D+\delta} [S_\delta(x + y - \delta) + C]\,dy. \end{aligned} \tag{1}$$

But

$$\lim_{x\to\infty} \int k(x - y)[S_\delta(y) + C]\,dy = \lim_{x\to\infty} \frac{1}{\delta} \int_x^{x+\delta} [g(y) + C]\,dy = C,$$

and it follows from Theorem A that for any positive D'

$$\lim_{x\to\infty} \frac{1}{D'} \int_0^{D'} [S_\delta(x + y) + C]\,dy = C.$$

Using this and (1), we obtain

$$\frac{(D - \delta)C}{D} \leqq \underline{\overline{\lim}}\, \frac{1}{D} \int_0^D [s(x + y) + C]\,dy \leqq \frac{(D + \delta)C}{D},$$

and since this is true for any positive δ,

$$\lim_{x\to\infty} \frac{1}{D} \int_0^D [s(x + y) + C]\,dy = C,$$

that is,

$$\lim_{x\to\infty} S_D(x) = 0. \tag{2}$$

It follows from (b) and the fact that

$$\int_x^{x+1} s(y)\,dy = S_1(x)$$

is bounded that

$$g^*(x) = \int k^*(x - y)s(y)\,dy$$

exists for all values of x. Let ϵ be > 0, and choose a positive integer N so that

$$\overline{\mathrm{bd}}\,|S_1(x)|\left\{\sum_{n=-\infty}^{-N-1} + \sum_{n=N}^{\infty}\ \overline{\mathrm{bd}}_{n \leqq x < n+1}\ |k^*(y)|\right\} < \epsilon.$$

Then

$$(3)\qquad \left|g^*(x) - \int_{-N}^{N} k^*(y)s(x-y)\,dy\right| = \left|\int_{-\infty}^{-N} + \int_{N}^{\infty} k^*(y)s(x-y)\,dy\right| < \epsilon.$$

Now since $k^*(x)$ is continuous almost everywhere, it is Riemann integrable in $(-N, N)$. We can therefore define two step functions $p(x)$, $q(x)$ such that

$$(4)\qquad p(y) \leqq k^*(y) \leqq q(y) \qquad (-N \leqq x \leqq N),$$

$$(5)\qquad \int_{-N}^{N} [q(y) - p(y)]\,dy \leqq \epsilon.$$

Then since $s(x) + C \geqq 0$,

$$\int_{-N}^{N} p(y)[s(x-y) + C]\,dy \leqq \int_{-N}^{N} k^*(y)[s(x-y) + C]\,dy$$
$$\leqq \int_{-N}^{N} q(y)[s(x-y) + C]\,dy.$$

But it follows immediately from (2), since $p(y)$, $q(y)$ are step functions, that

$$\lim_{x\to\infty} \int_{-N}^{N} p(y)[s(x-y) + C]\,dy = C\int_{-N}^{N} p(y)\,dy,$$
$$\lim_{x\to\infty} \int_{-N}^{N} q(y)[s(x-y) + C]\,dy = C\int_{-N}^{N} q(y)\,dy.$$

Hence

$$C\int_{-N}^{N} p(y)\,dy \leqq \underline{\overline{\lim}} \int_{-N}^{N} k^*(y)[s(x-y) + C]\,dy \leqq C\int_{-N}^{N} q(y)\,dy,$$

and using (4) and (5), we have

$$\overline{\lim_{x\to\infty}} \left|\int_{-N}^{N} k^*(y)s(x-y)\,dy\right| \leqq C\epsilon.$$

It follows from this and (3) that

$$\overline{\lim_{x\to\infty}}\,|g^*(x)| \leqq \epsilon(1 + C),$$

and since ϵ is arbitrary,

$$\lim_{x\to\infty} g^*(x) = 0.$$

HARVARD UNIVERSITY.

ORTHOGONAL POLYNOMIALS IN THREE VARIABLES

By Dunham Jackson

1. **Introduction.** The theory of orthogonal polynomials in two real variables naturally carries over to a considerable extent automatically to the case of three or more variables. For an adequate survey of the facts, nevertheless, it is necessary in some particulars to take explicit account of the greater complexity introduced by the increased number of dimensions. The Laplace series, for example, which can be regarded as an expansion in series of orthogonal polynomials on a sphere in space, is neither formally nor analytically a trivial extension of the Fourier series, which corresponds similarly to a circle in the plane.

This paper is introductory to the study of a class of developments among which the Laplace series is included as a special case, the primary aim being to clarify some of the new considerations that arise in making the transition from two variables[1] to a larger number. One question in particular relates to the determination of the number of polynomials of the n-th degree in the orthogonal system, a point which naturally has a dominant influence on the form of the resulting series developments. The discussion will be mainly for three variables, partly because the general outline of the extension to more than three will be sufficiently apparent, and partly, on the other hand, because a complete elucidation even of the three-dimensional case would be more than can be attempted here.

The range of integration with relation to which the property of orthogonality is defined may be a point set of more or less arbitrary character. It will be sufficient for purposes of illustration to take it as of finite extent, and to think of it as a region of space, a surface, or a curve. There will be occasion to distinguish between algebraic and non-algebraic surfaces, and in the case of curves to distinguish between algebraic curves, non-algebraic curves on an algebraic surface, and curves which do not lie on any algebraic surface. The curves and surfaces may or may not be closed. Algebraic curves and surfaces need not be complete algebraic loci; the surface considered may be the surface of a polyhedron, a hemispherical surface, or the complete surface of a solid hemisphere; the curve may be a skew polygon.

Occasion will be taken to refer to one matter, the uniqueness of the weight function for a given orthogonal system, which has not been discussed before even for the two-dimensional problem.

Received March 11, 1937; in revised form, March 4, 1938; presented to the American Mathematical Society, December 31, 1936, and November 27, 1937.

[1] See D. Jackson, *Formal properties of orthogonal polynomials in two variables*, this Journal, vol. 2(1936), pp. 423–434; *Orthogonal polynomials on a plane curve*, this Journal, vol. 3(1937), pp. 228–236.

2. **Three-dimensional region.** If V is a three-dimensional region and $\rho(x, y, z)$ a function which is non-negative and, for simplicity, almost everywhere positive on V, any finite number of the monomials

$$1, x, y, z, x^2, xy, xz, y^2, yz, z^2, x^3, \cdots \tag{1}$$

will be linearly independent on V, and the same will be true of the products of these monomials by $\rho^{\frac{1}{2}}$. The number of monomials of the n-th degree in the three variables jointly is $\frac{1}{2}(n+1)(n+2)$. Let ν, in conjunction with any particular value of n, stand for this number. By "Schmidt's process" it is possible to construct a system of normalized orthogonal polynomials $p_{nm}(x, y, z)$ $(n = 0, 1, 2, \cdots; m = 1, 2, \cdots, \nu)$, so that

$$\iiint_V \rho(x, y, z) p_{kl}(x, y, z) p_{nm}(x, y, z) dV = 0, \qquad |n - k| + |m - l| \neq 0,$$

$$\iiint_V \rho(x, y, z)[p_{nm}(x, y, z)]^2 dV = 1,$$

the first subscript in each case indicating the degree of the polynomial in the three variables together. As far as the individual polynomials are concerned, such a system can be formed in an infinite variety of ways; the system is determinate, however, to the extent that any set of ν polynomials of the n-th degree which are orthogonal to each other and to every polynomial of lower degree can be expressed in terms of any other such set by a ν-dimensional orthogonal transformation. The reasoning is an immediate adaptation of that in the case of two variables and need not be repeated. The same statement may be made with regard to the fact that if

$$\begin{aligned} x' &= A_{11}x + A_{12}y + A_{13}z, \\ y' &= A_{21}x + A_{22}y + A_{23}z, \\ z' &= A_{31}x + A_{32}y + A_{33}z \end{aligned} \tag{2}$$

is a linear transformation which carries over the region V and the weight function ρ into themselves, the ν polynomials $p_{nm}(x', y', z')$ for specified n are expressible by an orthogonal transformation in terms of the polynomials $p_{nm}(x, y, z)$.

3. **Non-algebraic surface.** Almost equally immediate is the setting up of a system of polynomials orthogonal on a non-algebraic surface, under convenient simplifying assumptions with regard to the data. The word *surface* will be understood throughout in a naïve geometric sense as referring to a closed point set of two-dimensional extent such that every one of its points is either an interior point two-dimensionally or a limit of such interior points. Let the surface S to be considered and the weight function ρ be furthermore supposed such that the integrations can be performed with respect to area on S. The hypothesis that the surface is not algebraic means that no linear relation of identity is satis-

fied on it by any finite number of the monomials (1). (Then there is no polynomial, other than zero, vanishing almost everywhere on S; for a polynomial vanishing almost everywhere on a "surface" as described would vanish by continuity at every point of the surface without exception.) The orthogonal system will contain $\frac{1}{2}(n+1)(n+2)$ polynomials of the n-th degree.

With respect to properties of symmetry and invariance under a linear transformation (2) of the variables, a completely general discussion would be complicated, in comparison with the problem in two dimensions, by questions of parametric representation. In simple cases, however, explicit consideration of such questions can be dispensed with. The essential point[2] is to recognize that if S and ρ are invariant under the transformation and if $P(x, y, z)$ is any polynomial in the variables, then

$$(3) \qquad \iint_S \rho(x, y, z)P(x', y', z')\,dS = \iint_S \rho(x, y, z)P(x, y, z)\,dS.$$

For reference later it is to be noted that as far as the present discussion of this particular point is concerned it is immaterial whether the surface is algebraic or not. If the transformation is a reflection in a plane, or a rotation through an angle $2\pi/k$, where k is an integer, with the effect in either case of permuting a finite number of congruent regions of S, the truth of (3) is apparent on consideration of each integral as limit of a sum, without recourse to any particular choice of parameters, the various regions being subdivided into congruent elements of area. If (3) holds under each of two transformations, it holds under their resultant. If S and ρ are invariant under an arbitrary rotation about a specified axis, the truth of (3) for a rotation of arbitrary magnitude about the axis follows by continuity from its validity for rotation through any rational multiple of π. Thus it is true here again, at least with a considerable degree of generality, that a linear transformation of the variables which leaves the range of integration and the weight function unchanged subjects the polynomials of the n-th degree in the orthogonal system to an orthogonal transformation.

4. **Algebraic surface.** In antithesis to the conditions of the preceding paragraphs let S be a surface on which some polynomial $\Omega_0(x, y, z)$ vanishes identically. If there is one such polynomial, there are of course infinitely many. In particular there is a polynomial $\Omega(x, y, z)$ vanishing identically on S, and composed of distinct irreducible factors each of which vanishes on a two-dimensional part of S. Any other polynomial $\Omega_1(x, y, z)$ which vanishes identically on S must be divisible by Ω.

In the sequence of monomials (1) let each be said for convenience to be of higher *rank* than those which precede it. The terms are ordered first with respect to degree in the three variables together, secondly with respect to degree in the pair of variables y, z, and thirdly with respect to degree in z. In any

[2] Cf. this Journal, vol. 3, loc. cit., p. 230.

polynomial let the term of highest rank with non-vanishing coefficient (or for brevity the corresponding monomial from (1), with the coefficient replaced by 1) be called the *leading term* of the polynomial. The leading term of the product of any two polynomials is the product of their leading terms.

Let the leading term of Ω be $x^p y^q z^r$, $p + q + r = N$. The leading term of any polynomial Ω_1 of the n-th degree which vanishes identically on S must be the product of $x^p y^q z^r$ by a monomial of degree $n - N$. The number of different monomial multipliers of this sort is $\frac{1}{2}(n - N + 1)(n - N + 2)$. So the number of monomials of the n-th degree, $n \geqq N$, which are *not* linearly dependent on monomials of lower rank on the surface S is

$$\tfrac{1}{2}(n + 1)(n + 2) - \tfrac{1}{2}(n - N + 1)(n - N + 2) = \tfrac{1}{2}N(2n + 3 - N).$$

It is possible therefore to construct a normalized orthogonal system containing this number of polynomials of the n-th degree for $n \geqq N$ (and $\frac{1}{2}(n + 1)(n + 2)$ polynomials of the n-th degree for $n < N$).

Let $\nu = \frac{1}{2}N(2n + 3 - N)$ or $\frac{1}{2}(n + 1)(n + 2)$ according to the value of n. Any set of ν normalized polynomials of the n-th degree which are orthogonal on S to each other and to every polynomial of lower degree is linearly expressible *on the surface* S in terms of any other such set by an orthogonal transformation.

As already mentioned, the discussion of the property of invariance expressed in (3), which determines the behavior of the orthogonal system under a transformation of (x, y, z) leaving S and ρ invariant, is independent of the algebraic or non-algebraic character of the surface; the resulting identities, however, now hold in general only on S.

As a case of special interest let S be the surface of the unit sphere, let the integrals be taken still with respect to area, and let the weight function be unity. Here $N = 2$, $\nu = 2n + 1$. As the $2n + 1$ harmonic homogeneous polynomials of the n-th degree which reduce on S to the elementary spherical harmonics have the requisite property of orthogonality, they may be taken, except for normalizing constant factors, as the orthogonal polynomials of the present theory. They are characterized here, however, as far as their values on the surface of the sphere are concerned, without explicit intervention of Laplace's equation. The partial sum of the Laplace series for a given function on the sphere is given by a polynomial chosen according to the least-square criterion among *all* polynomials of its degree.

5. **Identities.** Before we proceed to an examination of the relations that are found when the domain of integration is a curve, it may be appropriate to set down some observations which are applicable both in the cases that have been discussed and in those that remain to be considered.

Let the range of integration, whatever its dimensionality, be denoted by R. For each value of n let ν_n or ν be the number of polynomials of the n-th degree in the orthogonal system. Let these be denoted by $p_{n1}, \cdots, p_{n\nu}$. Any polynomial of the n-th degree can be expressed identically on R as a linear combina-

tion of the polynomials p_{ki} of degrees $\leqq n$. Let μ be the largest value of ν_k for $k \leqq n + 1$. For $\nu_k < i \leqq \mu$, and also for $k = -1$, let $p_{ki}(x, y, z)$ be defined as identically zero.

With this notation, whatever the manner of dependence of μ on n may be, a relation of recurrence, valid on R, can be written in the form[3]

$$xp_{ni} = \sum_j A_{nij}p_{n+1,j} + \sum_j B_{nij}p_{nj} + \sum_j C_{nij}p_{n-1,j},$$

if p_{ni} is any one of the polynomials $p_{n1}, \cdots, p_{n\mu}$, the sign $\sum_j$ indicating summation over j from 1 to μ. The recursion formula leads in the usual way to a Christoffel-Darboux identity for points (x, y, z) and (u, v, w) belonging to R: if

$$K_n \equiv K_n(x, y, z, u, v, w) = \sum_{k=0}^{n} \sum_{i=1}^{\mu} p_{ki}(x, y, z)p_{ki}(u, v, w),$$

then $(u - x)K_n$ can be written as

$$\sum_i \sum_j A_{nij}[p_{n+1,j}(u, v, w)p_{ni}(x, y, z) - p_{n+1,j}(x, y, z)p_{ni}(u, v, w)].$$

Replacement of $u - x$ by $v - y$ or $w - z$ affects this identity only to the extent of changing the values of the coefficients A_{nij}.

Since the right member of the Christoffel-Darboux identity as written consists of μ^2 pairs of terms (whether all different from zero or not), it is much more complicated in the case of a three-dimensional region, for which μ is of the order of magnitude of n^2, than on an algebraic surface, where μ is of the order of n, and is very greatly simplified, so as to be comparable with corresponding formulas in one variable, in cases to be discussed below in which μ is bounded for all values of n.

6. **Convergence.** Let the range of integration R be thought of as a closed point set contained in a cube K with edges parallel to the coördinate axes. If a continuous function f is given on R, and if this function is expanded in a series of the orthogonal polynomials corresponding to R with weight function unity, integration being performed with respect to volume, area, or arc length, convergence in the mean is readily deduced from the least-square property of the development. For the definition of f can be extended continuously[4] throughout K, and by Weierstrass' theorem f as thus extended can be uniformly approximated by means of polynomials throughout K and in particular on R.

The conclusion is almost equally immediate if unity as weight function is replaced by any integrable function having a positive lower bound. The same is true for a weight function having zero as greatest lower bound, if convergence in the mean is understood to refer to convergence of the weighted mean of the

[3] Cf. this Journal, vol. 3, loc. cit., pp. 235–236.

[4] See Hassler Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 63–89; especially p. 63.

square of the error toward zero. With a bounded weight function having a positive lower bound, the continuous function f can be replaced by an arbitrary function of class L^2, at least in the case of a curve (without double points) or a three-dimensional region; the corresponding problem for a surface appears to be complicated by questions of parametric representation. Convergence in the mean for a discontinuous f with an unbounded weight function naturally requires consideration of the properties of both functions jointly.

A similar method of approach is effective in certain cases in connection with the problem of uniform convergence. This is perhaps best illustrated in the case of approximation on the surface of the unit sphere, even though the outcome is in the first instance merely an alternative derivation of well known facts about Laplace series.

Let ρ, θ, φ be spherical coördinates, so that

$$x = \rho \cos \theta, \qquad y = \rho \sin \theta \cos \varphi, \qquad z = \rho \sin \theta \sin \varphi.$$

If $f(\theta, \varphi)$ is a function satisfying a Lipschitz condition with respect to arc length on the surface of the sphere, the function $F(x, y, z) = \rho f(\theta, \varphi)$ satisfies a Lipschitz condition in space and can be approximately represented[5] throughout the cube K by polynomials of the n-th degree with an error not exceeding a constant multiple of n^{-1}. Interpreted for $\rho = 1$ as a theorem on the approximate representation of $f(\theta, \varphi)$ by linear combinations of spherical harmonics,[6] this observation is of interest in itself. If $f(\theta, \varphi)$ has continuous directional derivatives on the spherical surface (i.e., continuous derivatives at each point with respect to latitude and longitude variables referred to a coördinate system for which the point in question is not a pole), $\rho^2 f(\theta, \varphi)$ has continuous partial derivatives with respect to x, y, and z and can be represented by polynomials of the n-th degree throughout K with a maximum error ϵ_n such that[7] $\lim_{n\to\infty} n\epsilon_n = 0$. Combined with an application of Bernstein's theorem, this yields a proof of convergence of the Laplace series[8] for $f(\theta, \varphi)$. Generalization is possible by introduction of a suitably qualified weight function. If the second directional derivatives of $f(\theta, \varphi)$ on the sphere are continuous, $\rho^3 f(\theta, \varphi)$ has continuous second partial derivatives as a function of (x, y, z); it can be approximated by polynomials with a maximum error approaching zero faster than n^{-2}. This

[5] D. Jackson, *Über die Genauigkeit der Annäherung stetiger Funktionen* · · · , Dissertation, Göttingen, 1911, pp. 88–92; see also the paper of Mickelson cited in footnote 7 below.

[6] For a proof based more explicitly on the properties of spherical harmonics see T. H. Gronwall, *On the degree of convergence of Laplace's series*, Transactions of the American Mathematical Society, vol. 15(1914), pp. 1–30; especially pp. 14–18.

[7] See E. L. Mickelson, *On the approximate representation of a function of two variables*, Transactions of the American Mathematical Society, vol. 33(1931), pp. 759–781; especially pp. 768–769. Mickelson points out in an introductory paragraph that the theorems on trigonometric and polynomial approximation which he presents at length for two variables can be extended to three or more.

[8] See Mickelson, loc. cit., pp. 776–780.

fact is useful in situations not previously treated which call for the use of Markoff's instead of Bernstein's theorem.

It is clear that the method is applicable in principle to problems of approximation on other domains than the spherical surface, though the precise formulation of results would require some attention to detail.

7. **Uniqueness of the weight function.** It may be pointed out in this connection that under conditions of considerable generality the weight function belonging to a given system of orthogonal polynomials is essentially determinate except for a constant factor. That is to say, if a complete system of polynomials (beginning with a constant as first term of the sequence) is orthogonal with respect to each of two weight functions, one of these functions is almost everywhere a constant multiple of the other.

Let the weight functions be ρ_1 and ρ_2, let the polynomials of the orthogonal system be denoted by p_{nm}, and let the domain of integration, whether volume, surface, or curve, be R. If $\int \rho_1 dR = k_1$, $\int \rho_2 dR = k_2$, and if the combination $\rho_1 - k_1 k_2^{-1} \rho_2$ is denoted by σ, then $\int \sigma dR = 0$. Since each non-constant p_{nm} is orthogonal to each of the other polynomials of the system, and in particular orthogonal to a constant, with respect to ρ_1 as weight function and also with respect to ρ_2,

$$\int \rho_1 p_{nm}\, dR = \int \rho_2 p_{nm}\, dR = 0, \qquad \int \sigma p_{nm}\, dR = 0.$$

So σ is orthogonal with unit weight function to every polynomial of the system, including the constant, or in other words is orthogonal on R to every polynomial. If σ is continuous, it can be represented on R by a uniformly convergent series of polynomials; being orthogonal to every term of the series, it is orthogonal to itself and identically zero. If σ is of class L^2, the domain R being taken as a volume or a curve as before to avoid complications of parametric representation, σ can be approximated in the mean by polynomials. If $\int \sigma^2 dR = I$, if ϵ is an arbitrary positive quantity, and if P is a polynomial such that $\int (\sigma - P)^2 dR < \epsilon/(I + 1)$, then by the orthogonality of σ to P and by Schwarz's inequality

$$\left[\int \sigma^2\, dR\right]^2 = \left[\int \sigma(\sigma - P)\, dR\right]^2 \leqq \int \sigma^2\, dR \int (\sigma - P)^2\, dR < \epsilon.$$

This means that $\int \sigma^2\, dR = 0$ and $\sigma = 0$ almost everywhere.

8. **Orthogonal polynomials on a curve.** Let attention be directed now more specifically to the formal properties of orthogonal polynomials on a space curve C, the curve being supposed rectifiable, and integrals taken with respect to arc length.

If each of the monomials (1) is linearly independent of the preceding ones on C, the orthogonal system again contains $\frac{1}{2}(n + 1)(n + 2)$ polynomials of the

n-th degree. The facts with regard to orthogonal transformation of the orthogonal polynomials, here and in subsequent cases, correspond to those previously noted.

When there is a polynomial that vanishes identically on C, it is of course not possible to say that all such polynomials must be multiples of a single one, as this would certainly not be true if C were the intersection of two algebraic surfaces. If there *is* a polynomial Ω of the N-th degree such that $\Omega = 0$ at all points of C and such that every polynomial vanishing identically on C is a multiple of Ω, construction of the orthogonal system proceeds as in the case of an algebraic surface, and the number of polynomials of the n-th degree in the system is again $\frac{1}{2}N(2n + 3 - N)$ for $n \geqq N$.

Particular interest attaches to the case of an algebraic curve, that designation being taken for present purposes to mean that there are two polynomials $\varphi(x, y, z)$, $\psi(x, y, z)$ which vanish identically on the curve, and which are relatively prime to each other. For it will be seen that for some classes of algebraic curves at least the number of polynomials of the n-th degree in the orthogonal system remains finite as n increases, and the theory of the corresponding expansions in series consequently may be expected to have an essentially closer resemblance to that of series of orthogonal polynomials in a single variable.

Algebraic curves, to be sure, do not constitute the only alternative to the cases noted above. Suppose, for example, that C consists of a segment of the z-axis and a non-algebraic curve in the xy-plane, both extending from the origin, say, so that jointly they form a single "curve" in the naïve sense. Every polynomial $\varphi(x, y, z)$ which vanishes on the whole of C contains z as a factor, since otherwise $\varphi(x, y, 0)$ would be a polynomial in x and y which does not vanish identically, but vanishes at all points of the non-algebraic curve. So no two polynomials vanishing on C can be relatively prime. On the other hand, z is the only factor common to all polynomials that vanish on C, since in particular xz and yz have no other common factor; but z itself does not vanish everywhere on C. It is clear from the earlier reasoning that if C is a curve on which a polynomial of the N-th degree vanishes identically, however this may be related to other polynomials having the same property, the number of polynomials of the n-th degree in the corresponding orthogonal system can not exceed $\frac{1}{2}N(2n + 3 - N)$ for $n \geqq N$. For the C just used as an illustration the number in question is $n + 2$ for $n \geqq 1$, the monomials of the n-th degree which are not linearly dependent on monomials of lower rank being those which are not divisible either by xz or by yz, i.e., $x^{n-k}y^k$ $(k = 0, 1, \cdots, n)$ and z^n.

The discussion of algebraic curves will be limited here to a few illustrative specifications. The general theory appears to lead to algebraic problems of some complexity. If the facts are well known from the elementary theory of space curves or the theory of algebraic functions, a formulation in terms appropriate to the problem in hand has not come to the writer's attention. The questions at issue do not particularly involve the properties of the systems of orthog-

onal polynomials as such, but are an essential preliminary to the study of such systems.

For any specified value of n let ν be the number of polynomials of the n-th degree in the orthogonal system. It is the number of monomials of the n-th degree which are on C linearly independent of monomials of lower rank. As in the situations previously considered, alternative sets of polynomials of the n-th degree in the orthogonal system are expressible in terms of each other on the range of integration by orthogonal transformation, and the value of ν is therefore independent of the particular notion of rank employed. Another observation which will be important is that ν can be defined as the number of polynomials of the n-th degree which on C are not connected with each other and with any polynomial of lower degree by any relation of linear dependence. That is to say, there exist ν polynomials $s_1(x, y, z), \cdots, s_\nu(x, y, z)$ of the n-th degree such that if $s(x, y, z)$ is any polynomial of degree lower than the n-th, a relation

$$a_1 s_1 + \cdots + a_\nu s_\nu + bs = 0$$

holding identically on C implies that $a_1 = \cdots = a_\nu = 0$, but if $s_1, \cdots, s_{\nu+1}$ are any $\nu + 1$ polynomials of the n-th degree, there exists a linear combination

$$a_1 s_1 + \cdots + a_{\nu+1} s_{\nu+1},$$

with coefficients not all zero, which reduces identically on C to a polynomial of lower degree (or to zero). As thus characterized, *ν is unaffected by any non-singular linear transformation of (x, y, z), whether orthogonal or not.* A permutation of the variables among themselves is of course a special case of such a transformation.

Suppose first that φ and ψ have leading terms $x^p y^q$ and z^r, respectively. Let $p + q = N_1$, and let r be denoted alternatively by N_2. Any monomial which is divisible by $x^p y^q$ or by z^r is linearly dependent on monomials of lower rank. If $x^h y^k z^l$ is not to be divisible by z^r, the exponent l must have one of the r values $0, 1, \cdots, r - 1$. For given $n = h + k + l$, the sum $h + k = n - l$ must have one of the r values $n, n - 1, \cdots, n - r + 1$. For a specific value of $h + k$, if $x^h y^k z^l$ is not to be divisible by $x^p y^q$, h must have one of the p values $0, 1, \cdots, p - 1$, or else one of the q values corresponding to $k = 0, 1, \cdots, q - 1$. So, when n is given, there are not more than r possibilities for the exponent l (fewer if $n < r - 1$), and with each of these at most $p + q$ possibilities for the exponent h, if $x^h y^k z^l$ is to be on C linearly independent of monomials of lower rank. Consequently

$$\nu \leqq (p + q)r = N_1 N_2,$$

the upper bound obtained for ν being the product of the degrees of φ and ψ. A corresponding conclusion is obtained if the leading terms are $x^p z^r$ and y^q, or x^p and $y^q z^r$, not by mere permutation of the variables in the statement of the conclusion itself (since the permutation would in general change the identity

of the leading terms),[9] but by repetition of the argument in detail. Also, one of the paired exponents may be zero, the leading terms being x^p and y^q, or x^p and z^r, or y^q and z^r.

Generally stated, the hypothesis of the last paragraph is that the leading terms of φ and ψ have no common factor. This of itself insures that φ and ψ are relatively prime, since the leading term of a common factor of φ and ψ would be a common factor of their leading terms. If the same variable occurs in both leading terms, and if no other hypothesis than specification of these leading terms is imposed, the argument necessarily fails, since the existence of a common factor of φ and ψ is not excluded; and if φ and ψ are not relatively prime, ν is not necessarily bounded as n increases, since a curve on which φ and ψ both vanish may be a non-algebraic curve on the surface defined by the vanishing of the common factor, as in an illustration already employed.

Nevertheless the boundedness of ν with increasing n can be deduced in some cases even when the leading terms of φ and ψ have a variable in common.

It will be assumed until the contrary is stated that the curve C is not plane, i.e., that there is no polynomial of the first degree in (x, y, z) which vanishes identically on C.

9. **Intersection of two quadrics.** Suppose, for example, that φ and ψ have leading terms x^2 and xy, respectively. If ψ were to contain a term in x^2, this could be removed by subtraction of a suitable multiple of φ, leaving a polynomial which must likewise vanish identically on C. For economy of notation it may be assumed without loss of generality that ψ lacks the term in x^2 in the first place. Then φ and ψ may be written in the form

$$\varphi(x, y, z) = x^2 + a_1 z + a_2 y + a_3 x + a_4,$$

$$\psi(x, y, z) = xy + b_1 z + b_2 y + b_3 x + b_4.$$

The combination $y\varphi - x\psi$, vanishing identically on C, has leading term yz if $a_1 \neq 0$, and has leading term y^2 if $a_1 = 0$, $a_2 \neq 0$. In either case φ and $y\varphi - x\psi$ are two polynomials of the second degree, vanishing identically on C, with leading terms relatively prime to each other, and it follows from the earlier reasoning that $\nu \leqq 4$. If $a_1 = a_2 = 0$, the polynomial φ does not contain y or z at all, and its vanishing represents a plane or a pair of parallel planes; it has been assumed that C does not lie in one plane, and as a connected curve it cannot be contained in two parallel planes without lying wholly in one of them. So it is certain under the hypotheses of the present paragraph that $\nu \leqq 4$.

[9] E.g., if the leading term of $\psi(x, y, z)$ is z^r, the leading term of $\psi(z, y, x)$ is not x^r in general; a polynomial with leading term x^r cannot have any other term of degree r, while a polynomial with leading term z^r may contain any or all of the monomials of the r-th degree. When a permutation of the variables is used below, in accordance with the last sentence of the preceding paragraph, all the terms of highest degree will be taken into account, not merely terms specified only by means of the notion of rank.

(If $a_1 = a_2 = 0$, if the equation $\varphi \equiv x^2 + a_3x + a_4 = 0$ has distinct real roots α, β, and if the word *curve* is used in a less restrictive sense to admit the possibility of a locus C consisting of separate curves in the planes $x = \alpha$, $x = \beta$, respectively, the product $\psi(\alpha, y, z)\psi(\beta, y, z)$ is a polynomial of the second degree (at most) vanishing identically on C, with leading term z^2 if $b_1 \neq 0$, y^2 if $b_1 = 0$, $(b_2 + \alpha)(b_2 + \beta) \neq 0$, so that $\nu \leqq 4$ again; if $b_1 = 0$, $b_2 = -\alpha$, say (it is immaterial which root is denoted by α), then

$$\psi = (x - \alpha)y + b_3(x - \alpha) + b_3\alpha + b_4\,,$$

which is inconsistent with the hypotheses, giving ψ and φ the common factor $x - \alpha$ if $b_3\alpha + b_4 = 0$, and making ψ different from zero at all points of the plane $x = \alpha$ if $b_3\alpha + b_4 \neq 0$, so that C, on which ψ vanishes identically, is restricted to the single plane $x = \beta$.)

For reference a few lines below it is to be noted that *if three* polynomials with leading terms xy, xz, yz vanish identically on C, then $\nu \leqq 3$. For every monomial of the n-th degree other than x^n, y^n, z^n is divisible by one (or more) of the three leading terms, and so is linearly dependent on monomials of lower rank.

Suppose next that φ and ψ have leading terms xy and xz:

$$\varphi(x, y, z) = xy + a_0x^2 + a_1z + a_2y + a_3x + a_4\,,$$

$$\psi(x, y, z) = xz + b_0x^2 + b_1z + b_2y + b_3x + b_4\,.$$

There is no loss of generality in omitting xy from ψ, since it could be removed by subtraction of a multiple of φ. The non-singular substitution

$$x' = x, \qquad y' = y + a_0x, \qquad z' = z + b_0x$$

gives polynomials of the same form in x', y', z', except that terms in x'^2 are lacking; instead of working with the new notation it may be assumed with essential generality at the outset that φ, ψ have the form

$$\varphi = xy + a_1z + a_2y + a_3x + a_4\,,$$

$$\psi = xz + b_1z + b_2y + b_3x + b_4\,.$$

The polynomial $\omega = z\varphi - y\psi$ is of the form

$$a_1z^2 + (a_2 - b_1)yz - b_2y^2 + \text{terms of lower rank than } y^2.$$

If $a_1 \neq 0$, ω has leading term z^2, and the vanishing of φ and ω on C means that $\nu \leqq 4$. If $a_1 = 0$, $a_2 - b_1 \neq 0$, the leading term of ω is yz, and the simultaneous vanishing of φ, ψ, and ω means that $\nu \leqq 3$. If $a_1 = 0$, $a_2 - b_1 = 0$, $b_2 \neq 0$, the leading term is y^2, and ψ and ω together give $\nu \leqq 4$. If $a_1 = a_2 - b_1 = b_2 = 0$, φ and ψ reduce to

$$\varphi = xy + a_2y + a_3x + a_4\,,$$

$$\psi = xz + a_2z + b_3x + b_4\,.$$

Then

$$(z + b_3)\varphi - (y + a_3)\psi = (a_4 - a_2a_3)z - (b_4 - a_2b_3)y + a_4b_3 - a_3b_4 .$$

The vanishing of this expression on C would mean that C is a plane curve, unless the right member is zero identically, in which case φ and ψ have a common factor.[10] Both of these suppositions having been ruled out, the coefficients must come under one of the hypotheses which make $\nu \leqq 4$.

Now let φ and ψ be any two polynomials of the second degree which are relatively prime. By a non-singular linear transformation it can be insured that φ contains a term in z^2, which is then its leading term. If ψ contains a term in z^2, this can be removed by subtracting a multiple of φ. If the resulting polynomial contains neither yz nor xz, the leading term being y^2, xy, or x^2, this leading term has no factor in common with the leading term of φ, and $\nu \leqq 4$. If the polynomial contains yz, this is the leading term; if it contains xz but not yz, interchange of x and y makes yz the leading term again. So the cases requiring further examination can be covered by the assumption that φ and ψ have z^2 and yz as leading terms. The term yz furthermore can be removed from φ by subtracting a multiple of ψ.

Let the polynomials then be taken as

$$\varphi(x, y, z) = z^2 + a_1y^2 + a_2xz + a_3xy + a_4x^2 + \cdots ,$$

$$\psi(x, y, z) = yz + b_1y^2 + b_2xz + b_3xy + b_4x^2 + \cdots .$$

The continuation sign $+ \cdots$ will henceforth be understood always to mean "plus terms of lower degree". The combination $y\varphi + (b_1y - z)\psi$ has leading term y^3, if $a_1 + b_1^2 \neq 0$. The vanishing of this polynomial on C together with φ implies that $\nu \leqq 6$. Specifically, this is so because all monomials of the n-th degree except

$$x^n, \qquad x^{n-1}y, \qquad x^{n-2}y^2, \qquad x^{n-1}z, \qquad x^{n-2}yz, \qquad x^{n-3}y^2z$$

are divisible either by z^2 or by y^3. But of these six the last two are divisible by yz, the leading term of ψ, and so as functions on C these also are linearly dependent on monomials of lower rank. The relation $\nu \leqq 6$ is thus replaced[11] by $\nu \leqq 4$.

If $a_1 + b_1^2 = 0$,

$$\varphi(x, y, z) = z^2 - b_1^2y^2 + a_2xz + a_3xy + a_4x^2 + \cdots ,$$

while ψ remains as before. Let the variables be subjected to the transformation

$$x_1 = x, \qquad y_1 = y, \qquad z_1 = z + b_1y.$$

[10] Elimination of xy from ψ has, to be sure, replaced the most general φ, ψ by φ and $\psi - k\varphi$, where k is a constant. But if φ and $\psi - k\varphi$ have a common factor, the same is of course true of φ and ψ.

[11] In the preceding discussion for the cases of leading terms x^2, xy, and xy, xz, the conclusion $\nu \leqq 4$ can similarly be replaced by $\nu \leqq 3$.

This makes

$$\varphi(x, y, z) = \varphi_1(x_1, y_1, z_1)$$
$$= z_1^2 - 2b_1y_1z_1 + a_2x_1z_1 + (a_3 - a_2b_1)x_1y_1 + a_4x_1^2 + \cdots,$$
$$\psi(x, y, z) = \psi_1(x_1, y_1, z_1) = \quad y_1z_1 + b_2x_1z_1 + (b_3 - b_1b_2)x_1y_1 + b_4x_1^2 + \cdots.$$

The combination

$$\omega(x, y, z) = \varphi(x, y, z) + 2b_1\psi(x, y, z)$$
$$= \varphi_1(x_1, y_1, z_1) + 2b_1\psi_1(x_1, y_1, z_1) = \omega_1(x_1, y_1, z_1)$$

has the form

$$\omega_1(x_1, y_1, z_1) = z_1^2 + a_{12}x_1z_1 + a_{13}x_1y_1 + a_{14}x_1^2 + \cdots,$$

in which the manner of dependence of the new coefficients on the old ones need not be further specified. For the sake of uniformity let ψ_1 be written as

$$\psi_1(x_1, y_1, z_1) = y_1z_1 + b_{12}x_1z_1 + b_{13}x_1y_1 + b_{14}x_1^2 + \cdots.$$

The original φ and ψ having been replaced by ω_1 and ψ_1, the condition $a_1 + b_1^2 = 0$ is seen to be essentially equivalent to the vanishing of a_1 and b_1 separately.

Let

$$x_2 = x_1, \qquad y_2 = y_1 + b_{12}x_1, \qquad z_2 = z_1 + \tfrac{1}{2}a_{12}x_1$$

This transformation has the effect of removing the terms in x_1z_1, and reduces the polynomials to the form

$$\omega_1(x_1, y_1, z_1) = \omega_2(x_2, y_2, z_2) = z_2^2 + a_{23}x_2y_2 + a_{24}x_2^2 + \cdots,$$
$$\psi_1(x_1, y_1, z_1) = \psi_2(x_2, y_2, z_2) = y_2z_2 + b_{23}x_2y_2 + b_{24}x_2^2 + \cdots.$$

Here let

$$\xi = x_2 + \lambda y_2, \qquad \eta = y_2, \qquad \zeta = z_2,$$

with an arbitrary λ. In terms of the new variables, ω_2, ψ_2 have the form

$$\omega_2(x_2, y_2, z_2) = \Omega(\xi, \eta, \zeta) = \zeta^2 + \alpha_1\eta^2 + \alpha_3\xi\eta + \alpha_4\xi^2 + \cdots,$$
$$\psi_2(x_2, y_2, z_2) = \Psi(\xi, \eta, \zeta) = \eta\zeta + \beta_1\eta^2 + \beta_3\xi\eta + \beta_4\xi^2 + \cdots,$$

with $\alpha_1 = a_{24}\lambda^2 - a_{23}\lambda$, $\beta_1 = b_{24}\lambda^2 - b_{23}\lambda$. These polynomials have the form of the original φ and ψ (with the specialization that the coefficients of $\xi\zeta$ are zero). If $\alpha_1 + \beta_1^2 \neq 0$, repetition of the earlier argument shows that $\nu \leqq 4$. It will be possible to choose λ so as to make $\alpha_1 + \beta_1^2 \neq 0$, unless

$$a_{24}\lambda^2 - a_{23}\lambda + (b_{24}\lambda^2 - b_{23}\lambda)^2 = 0$$

identically in λ, or

$$a_{23} = 0, \qquad b_{24} = 0, \qquad a_{24} + b_{23}^2 = 0.$$

The discussion has to be resumed only under the hypothesis that these special conditions are fulfilled.

The polynomials to be considered are then

$$\omega_2(x_2, y_2, z_2) = z_2^2 - b_{23}^2 x_2^2 + \cdots,$$

$$\psi_2(x_2, y_2, z_2) = y_2 z_2 + b_{23} x_2 y_2 + \cdots.$$

Let

$$x_3 = z_2 + b_{23} x_2, \qquad y_3 = y_2, \qquad z_3 = x_2.$$

The polynomials become

$$\omega_2(x_2, y_2, z_2) = \omega_3(x_3, y_3, z_3) = x_3^2 - 2b_{23} x_3 z_3 + \cdots,$$

$$\psi_2(x_2, y_2, z_2) = \psi_3(x_3, y_3, z_3) = x_3 y_3 + \cdots.$$

These have leading terms $x_3 z_3$ and $x_3 y_3$, unless $b_{23} = 0$, in which case the leading terms are x_3^2 and $x_3 y_3$. Both cases have been treated in earlier paragraphs, with the conclusion each time that $\nu \leqq 4$. This relation for ν is therefore of general validity.

It has been assumed for definiteness that C is not a plane curve. If it does lie in a plane, this can be reduced by linear transformation to the plane $z = 0$. The whole problem then is one of polynomials in the two variables x, y. If the coördinates of all points of C satisfy an equation of the N-th degree,[12] $\nu \leqq N$. The special case of a plane curve need not therefore be excluded from the following general statement:

If two polynomials of the second degree which are relatively prime vanish identically on C, the number of polynomials of the n-th degree in the corresponding orthogonal system does not exceed 4 for any value of n.

THE UNIVERSITY OF MINNESOTA.

[12] See the second of the papers cited in footnote 1.

A LATTICE FORMULATION FOR TRANSCENDENCE DEGREES AND p-BASES

By Saunders Mac Lane

1. **Introduction**. The transcendence degree of an extension of a field is the cardinal number of a maximal set of independent transcendents in the extension (Steinitz [11]);[1] in an Abelian group without elements of finite order the *rank* is the cardinal number of a maximal set of rationally independent group elements (Baer [1]). Both these cardinal numbers are invariants; the proofs of these two facts are similar, so that there should be an underlying theorem generalizing these proofs and stated in terms of the lattice of subfields or of subgroups, as the case may be.[2] This paper constructs, in §2, a type of lattice, called an "exchange" lattice, in which such a theorem can be proved (cf. §3). This lattice theorem includes also some investigations of Teichmüller ([12] and [13]) on fields of characteristic p; in particular, we establish the invariance of the cardinal number of a "relative p-basis" of an inseparable algebraic extension of such a field.

The crucial axiom for our lattices is an "exchange" axiom, related to the Steinitz exchange theorem. This axiom is equivalent to one of the axioms recently used by Menger [7] in investigating the algebra of affine geometry, and also to a certain covering property used by Birkhoff [2] in an analysis of the Jordan Theorem (cf. §4). This axiom can be viewed as a weakened form of the Dedekind or modular axiom for a lattice (Ore [9], or Birkhoff [4]). The Dedekind axiom itself could not apply to the lattices of fields with which we are concerned (see §5 and §6).

Unfortunately the exchange axiom is stated in terms of the "points" of the lattice, or alternately in terms of the covering relation. It thus applies only trivially to continuous geometries or to other infinite lattices having no points. In §7 we succeed in constructing two exchange axioms which in the presence of the other axioms are equivalent to the original exchange axiom, but which themselves do not involve points or coverings. These new axioms yield most of the usual properties of the dimension function in a finite lattice. Their title to be considered as substitutes for the modular law rests chiefly on their versatility: each of the new exchange axioms is seen in §8 to be equivalent to a natural assertion about the possibility of specified types of transpositions in any given chain of the lattice.

Received March 1, 1938; presented to the American Mathematical Society, April 9, 1938.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] This remark is due to R. Baer (in conversation).

2. **Exchange lattices**. The transcendence degree of a field $\mathfrak{L}$ over a subfield $\mathfrak{K}$ can be defined in terms of those subfields $\mathfrak{M}, \mathfrak{K} \subset \mathfrak{M} \subset \mathfrak{L}$, which are relatively algebraically closed in $\mathfrak{L}$, in the sense that every element of $\mathfrak{L}$ algebraic over $\mathfrak{M}$ is contained in $\mathfrak{M}$. With respect to field inclusion, these subfields $\mathfrak{M}$ form a continuous lattice.

A set L is a *continuous lattice*[3] (in the terminology of O. Ore, a *complete structure*) if a transitive and irreflexive relation $a < b$ is so defined for elements a and b of L that for every subset $A \subset L$ there exist in L an element $\Sigma(A)$, the *union*, and an element $\Pi(A)$, the *cross-cut*, such that $c \geqq \Sigma(A)$ holds if and only if $c \geqq a$ for every a in A, while $d \leqq \Pi(A)$ if and only if $d \leqq a$ for every a in A. If A consists of two elements a and b, we denote the union or join by $a + b$, and the cross-cut or meet by $a \cdot b$. The cross-cut $0 = \Pi(L)$ is the *zero* element, the union $1 = \Sigma(L)$ the *unit* element of L. If $a > b$ in L, but $a > c > b$ is impossible for c in L, then a is said to *cover* b or to be prime over b (Ore [9]). An element p prime over 0 is called a *point* of L (Menger [7]).

In the field case, each indeterminate x over $\mathfrak{K}$ generates a relatively algebraically closed field $\mathfrak{K}(x)'$ which is a point in the lattice of fields $\mathfrak{M}$. The crucial property of algebraic dependence is: if x is algebraic over $\mathfrak{M}(y)$, but not over $\mathfrak{M}$, then y is algebraic over $\mathfrak{M}(x)$. For lattices, we state a corresponding *exchange axiom*:

(E_1) If a is in L while p and q are points of L, then $a < a + p \leqq a + q$ implies $q \leqq a + p$.

We need also the existence of points and the "finiteness" of dependence:

(G_1) If $b < a$ in L, then there is in L a point p with $b < b + p \leqq a$.

(F_1) If Q is a set of points and p a point of L with $p \leqq \Sigma(Q)$, then there exists a finite set of points $q_1, q_2, \cdots, q_n$ of Q with $p \leqq q_1 + q_2 + \cdots + q_n$.

Definition. *An exchange lattice is a continuous lattice satisfying* (E_1), (G_1), *and* (F_1).

Note that (E_1) is equivalent to the assertion that $a < a + p < a + q$ is impossible, while (G_1) is equivalent to the following formally stronger axiom:

(G_2) If $a \nleqq b$, then there is a point p with $p \leqq a$, $p \nleqq b$.

For $a \nleqq b$ implies $ab < a$, hence by (G_1) there is a p such that $p \leqq a$, $p \nleqq ab$. Thus $p \nleqq b$. In the presence of (G_1), (F_1) can be shown equivalent to the following point-free statement:

(F_2) If $b \leqq \Sigma(A)$, $A \subset L$, then there are in A elements $a_1, \cdots, a_n$ with $b(a_1 + \cdots + a_n) > 0$.

To prove (F_2), let Q be the set of all points $q \leqq$ some a of A, and choose any $p \leqq b$. Then $p \leqq \Sigma(Q) = \Sigma(A)$ by (G_2), whence $p \leqq q_1 + \cdots + q_n$ by (F_1). This gives the conclusion.

Our postulates are related to the properties of algebraic dependence as formulated by van der Waerden [14]. He considers a relation "b depends on S" for an element b and a subset S of a given set D, with the following properties:

[3] This definition is in von Neumann [8]. For further literature on lattices, cf. Ore [10] or Köthe [5].

1. each b depends on the set $\{b\}$;
2. if b depends on S and $S \subset T$, then b depends on T;
3. if b depends on S, then b depends on some finite subset of S;
4. if b depends on $S_0 = \{c_1, c_2, \cdots, c_n\}$ but on no proper subset of S_0, then c_n depends on $\{c_1, \cdots, c_{n-1}, b\}$;
5. if b depends on S and if every element of S depends on T, then b depends on T.

A relation "a depends on S" with these five properties we call for the moment a *dependence relation.* The exact connection with our axioms we state without proof as follows:

THEOREM 1. *If D is a set with a dependence relation, and if a subset $S \subset D$ is said to be closed whenever S contains every b of D dependent on S, then the set $\mathfrak{L}(D)$ of all closed subsets of D is an exchange lattice if $S < T$ means that S is a proper subset of T. Conversely, if L is any exchange lattice, and if $D = \mathfrak{D}(L)$ is the set of all points of L, then the relation $p \leqq \Sigma(S)$ for $S \subset D$ is a dependence relation of p to S. Furthermore, $\mathfrak{L}(\mathfrak{D}(L))$ is isomorphic to the given lattice L, while, for a given D, the set $D' = \mathfrak{D}(\mathfrak{L}(D))$ with its dependence relation is isomorphic to the set D^* obtained from D by identifying all pairs of mutually dependent elements of D. Here an isomorphism of D' to D^* means a one-to-one correspondence of D' to D^* which leaves unchanged the dependence relation.*

In the finite case, there is a similar connection to the matroids of Whitney [15], as expressed in terms of lattices by Birkhoff [3].

3. **The basis theorem**.

A transcendence basis for a field is a maximal set of algebraically independent elements. Generally, a set P of points in any lattice L is *independent* if $(\Sigma(P')) \cdot (\Sigma(P'')) = 0$ for any two disjoint subsets P' and P'' of P (von Neumann [8], Chapter II). An independent set P of points with union $\Sigma(P) = a$ is called a *basis* of the element a of L. In the transcendence basis case, an independent set corresponds exactly to a set of indeterminates irreducible in the sense that no one indeterminate depends algebraically on the others. The same definition of independence holds in other cases:

THEOREM 2. *A set P of points in an exchange lattice is independent if and only if $p \leqq p_1 + p_2 + \cdots + p_n$ is impossible for distinct points $p, p_1, \cdots, p_n$ of P. Hence a set P is independent if and only if every finite subset of P is independent.*

Proof. The necessity of the condition is immediate. Suppose conversely that $p \leqq \Sigma p_i$ is impossible, but that P is dependent. There then are disjoint subsets P' and $P'' \subset P$ with $(\Sigma(P')) \cdot (\Sigma(P'')) > 0$. There then exists a point q such that $0 < q \leqq (\Sigma(P')) \cdot (\Sigma(P''))$. By the finiteness axiom, points p_i' in P' and p_j'' in P'' can be chosen so that

$$q \leqq p_1' + p_2' + \cdots + p_m', \qquad q \leqq p_1'' + p_2'' + \cdots + p_n''.$$

If $m > 0$ is so small that $q \not\leqq p_1' + \cdots + p_{m-1}'$, the exchange axiom implies that

$$p_m' \leqq p_1' + \cdots + p_{m-1}' + q \leqq p_1' + \cdots + p_{m-1}' + p_1'' + \cdots + p_n'',$$

contrary to the assumed property of P. Hence P is independent. A similar use of the exchange axiom gives also the result (cf. Menger [7], p. 462):

COROLLARY. *If P is an independent set in the exchange lattice L, and if q is a point with $q(\Sigma(P)) = 0$, then the set obtained by adjoining q to P is independent.*

THEOREM 3. *If L is an exchange lattice and if P is any independent set of points of L with $\Sigma(P) \leqq b$, then there is a set $Q \supset P$ which is a basis of b. In particular, every element of L has a basis.*

Proof. When $b = 1$, we can construct Q from P by well ordering the points of L and by applying repeatedly Theorem 2 and its corollary. A basis for any $b \neq 1$ can then be found by applying the previous case to the quotient lattice $b/0$ (Ore [9], p. 425). For any $b \geqq c$ such a *quotient lattice* b/c consists of all elements d with $b \geqq d \geqq c$.

THEOREM 4. *If $b \geqq c$ are elements of the exchange lattice L, then the quotient lattice b/c is also an exchange lattice. Its points are all the elements of L of the form $c + p$, where p is a point of L such that $c < c + p \leqq b$.*

Proof. Any such $c + p$ is a point of b/c, for otherwise $c + p > d > c$, hence by (G_1) $c + p > c + q > c$, and this is impossible by the exchange axiom. Conversely, the existence axiom for points of L shows that any point of b/c must have the form $c + p$. The axioms for an exchange lattice can now be directly verified for b/c.

The comparison of different bases depends upon an exchange process (Steinitz [11], Theorem 7, p. 115).

THEOREM 5. *If P is an independent set of points in an exchange lattice, q a point with $q \leqq \Sigma(P)$, then there is a finite subset $P_0 \subset P$ such that* (i) $q \leqq \Sigma(P_0)$, *but this statement is false if P_0 is replaced by any of its proper subsets*; (ii) *if $R \subset P$ and $q \leqq \Sigma(R)$, then $P_0 \subset R$*; (iii) *if in P any point p_1 of P_0 be replaced by q, the resulting set P' is independent, and $\Sigma(P') = \Sigma(P)$.*

Proof. The finiteness axiom yields at once the set P_0 as in (i). If $q \leqq \Sigma(R)$, as in (ii), then $q \leqq \Sigma(R_0)$ for a similar minimal finite set R_0. If $P_0 \subset R$ is false, then $P_0 \neq R_0$. Since P_0 is minimal, R_0 contains at least one point r not in P_0. If R_0' is the set R_0 with r deleted, then $q \leqq \Sigma(R_0') + r$, so that the exchange axiom proves $r \leqq \Sigma(R_0') + q \leqq \Sigma(R_0') + \Sigma(P_0)$, contrary to the independence of P. Finally P' in conclusion (iii) is independent because of (ii) and the corollary to Theorem 2.

Repeated applications of this exchange process with a transfinite induction yield, exactly as in Steinitz [11] (Note 126 and correction thereto), a proof of the fundamental invariance theorem:

THEOREM 6. *If P and Q are two bases for an element b in an exchange lattice L, then the sets P and Q have the same cardinal number.*[4] *This number we call the rank of b.*

THEOREM 7. *Every exchange lattice L has relative complements; that is, given $b \geqq d \geqq c$ in L there is an element d' in L with $d + d' = b$, $dd' = c$.*

[4] An equivalent abstract basis theorem has been developed by Reinhold Baer, who also found several alternative postulational bases for abstract independence (all unpublished). (Added May 9, 1938.)

Proof. In the quotient lattice b/c there is by Theorems 4 and 3 a basis P_d for d. By Theorem 3 we can adjoin to P_d a set of points Q' of b/c such that P_d and Q' together form a basis for b. Then $d' = \Sigma(Q')$ is the desired complement, because the definition of independence insures that $(\Sigma(P_d))\cdot(\Sigma(Q')) = c$.

4. **Alternative exchange axioms and the Dedekind axiom.** The exchange axiom is a weakened form of the usual modular or Dedekind law (Ore [9], p. 412; Birkhoff [4]).

(D_1) DEDEKIND AXIOM. $a \geqq c$ implies $a(b + c) = ab + c$.

This axiom holds for the linear subspaces of a projective space, but not for an affine space. Menger showed that it could there be replaced in part by either of the two equivalent assertions, for all elements a, b and all points p of L, that

(E_2) $a \leqq b \leqq a + p$ implies $b = a$ or $b = a + p$; or that

(E_3) $p \nleqq a + b$ implies $ab = (a + p)b$

(see [7], Axiom 6^+). Birkhoff's investigations of the Jordan Theorem[5] involve the property that, for all a, b and d in L,

(E_4) if $a \geqq d$, $a + b > a$ and b covers d, then $a + b$ covers a.

The interrelations of these axioms are as follows.

THEOREM 8. *In any lattice L satisfying the point-existence axiom* (G_1), *any two of the conditions* (E_1), (E_2), (E_3), (E_4) *are equivalent. Each of them is a consequence of the Dedekind axiom, but there exist exchange lattices which satisfy* (E_1), (F_1), *and* (G_1) *but not the Dedekind axiom.*

In an arbitrary lattice the implications (E_4) $\rightarrow$ (E_2) $\rightarrow$ (E_1), (E_2) $\leftrightarrow$ (E_3) (proved by Menger) and (D_1) $\rightarrow$ (E_4) (cf. Birkhoff [2], Theorem 9.1) will hold. To prove (E_4) $\rightarrow$ (E_2), set $d = 0$ and $b = p$ in (E_4) to obtain the assertion "$a + p > a$ implies that $a + p$ covers a", an alternative statement for (E_2). Similarly (E_2) with $b = a + q$ will yield (E_1).

To establish (E_4) as a consequence of (E_1) and (G_1), suppose that in the conclusion of (E_4) $a + b$ fails to cover a. There is then a c, $a + b > c > a$, and by (G_1) there are points p and q such that $a < a + p \leqq c$, $d < d + q \leqq b$. Since b covers d, $d + q = b$ and

$$a < a + p \leqq c < a + b = a + d + q = a + q.$$

The exchange axiom then makes $a + q \leqq a + p < a + q$, a contradiction. Given (G_1) we now have (D_1) $\rightarrow$ (E_4) $\rightarrow$ (E_3) $\rightarrow$ (E_2) $\rightarrow$ (E_1) $\rightarrow$ (E_4), as asserted in the theorem.

The Dedekind axiom does not hold in every exchange lattice. For instance, let C be a finite class with m elements and n an integer, $1 \leqq n < m$. Consider the set L_{mn} of all subclasses $a, b, \cdots$ of C which have not more than n elements or exactly m elements, and let $a < b$ mean that a is a proper subclass of b.

[5] Lattices having this property (E_4) have been studied by F. Klein, who calls the property "axiom (γ'')" and who obtains several interesting equivalent axioms. Cf. F. Klein, *Birkhoffsche und harmonische Verbände*, Mathematische Zeitschrift, vol. 42 (1937), pp. 58–81. (Added May 9, 1938.)

Then L_{mn} is a lattice, in which the points are the subclasses of C with just one element, while the sum $a + b$ is the whole class C or the ordinary point-set union $a \cup b$ according as $a \cup b$ has more than n elements or not. The axioms for an exchange lattice are readily verified for L_{mn}. Whenever $m \geqq n + 2 \geqq 4$, and in no other cases, L_{mn} fails to be a Dedekind lattice. For instance, if $m = 4$, $n = 2$, L_{42} contains four points p, q, r and s. If we set $a = p + q$, $b = r + s$, $c = p$, then $a \geqq c$, but

$$a(b + c) = (p + q)(p + r + s) = (p + q)\cdot 1 = p + q,$$

$$ab + c = (p + q)(r + s) + p = p \text{ and } p + q \neq p,$$

contrary to the Dedekind law.

We turn to the Jordan Theorem. A *chain* of length n joining a to b, $a < b$, is a sequence of $n + 1$ elements x_i, $a = x_0 < x_1 < \cdots < x_n = b$. The chain is *principal* if each x_i covers x_{i-1}, for $i = 1, \cdots, n$. A lattice L is said to be of *finite dimensions* if any two elements $a < b$ can be joined by a principal chain.

THEOREM 9. *If two elements $a < b$ of an exchange lattice L are joined by a principal chain of length n, then no chain joining a to b has length greater than n, so that any principal chain joining a to b has the same length n. If the exchange lattice L is of finite dimensions, then the rank (or dimension) $\rho(a)$ defined in Theorem 6 is equal to the length of any principal chain joining 0 to a, and $\rho(a)$ satisfies the inequality*

$$\rho(ab) + \rho(a + b) \leqq \rho(a) + \rho(b). \tag{1}$$

The first assertion, which states in part that when a and b are joined by a principal chain they can be joined by no "infinite" chain, can be established by induction, with repeated applications of axiom (E_4). The rest of the theorem restates known results, for the exchange axiom (E_4) implies the following condition:

(ξ) If a and b cover d and $a \neq b$, then $a + b$ covers a and b. From this condition the Jordan Theorem and (1) can be established (Birkhoff [2], Theorems 8.2 and 9.2; Ore [9], Chapter II).

The following converse to Theorems 7 and 9 gives in effect an alternative definition for finite exchange lattices.

THEOREM 10. *If L is a lattice which has for every $a > d > 0$ a "relative complement" d' with $d + d' = a$, $d \cdot d' = 0$, and if to every element a of L an integer $\rho(a)$ can be so assigned that*

(i) *$a < b$ implies $\rho(a) < \rho(b)$;*

(ii) *if b covers a, then $\rho(b) = \rho(a) + 1$;*

(iii) *the inequality (1) holds for all a and b;*

then L is an exchange lattice.

Proof. The finiteness of each $\rho(a)$ and the condition (i) enforce ascending and descending chain conditions (Ore [9], p. 410) and thence the existence of infinite unions and cross-cuts, as well as the finiteness axiom (F_1). To establish the existence of points, let $a > d \geqq 0$ and pick a relative complement d' of d in a.

The descending chain condition yields a point p, $p \leqq d'$, with the requisite properties, $p \leqq a$ and $p \nleqq d$.

If p is any point with $a < a + p$, then the inequality (1) and condition (ii) give

$$\rho(a + p) \leqq \rho(a) + \rho(p) - \rho(ap) = \rho(a) + \rho(p) - \rho(0) = \rho(a) + 1.$$

But $a + p > a$, so $\rho(a + p) > \rho(a)$ by (i), and we have $\rho(a + p) = \rho(a) + 1$. This means that $a + p$ covers a, and is in effect the exchange axiom (E_2). Therefore L is an exchange lattice.

A lattice L is *complemented* if for every a there is an a' with $a + a' = 1$, $aa' = 0$. If L is also modular, relative complements are known to exist, and in the finite case there is a dimension function. Therefore, we have the following

COROLLARY. *Any complemented Dedekind lattice of finite dimensions is an exchange lattice.*

The lattice of all linear subspaces of an affine space is also an exchange lattice, as can be established from Menger's axioms for affine geometry.

5. **Transcendence degrees of fields.** For fields $\mathfrak{L} \supset \mathfrak{K}$, the set of all relatively algebraically closed subfields $\mathfrak{M}$ between $\mathfrak{K}$ and $\mathfrak{L}$ forms an exchange lattice L. This is proved by noting that the relation "y depends algebraically on $\mathfrak{K}(S)$" is a dependence relation with the five properties used in Theorem 1 to construct an exchange lattice. The sets "closed" under this relation are exactly the subfields $\mathfrak{M}$, and the transcendence degree of $\mathfrak{M}$ over $\mathfrak{K}$ is the rank of $\mathfrak{M}$ in the lattice, and so is included in Theorem 6.

Such lattices of fields need not satisfy the Dedekind law. Consider over any field $\mathfrak{K}$ the field $\mathfrak{L} = \mathfrak{K}(x, y, z)$ of rational functions of three independent variables x, y and z. The subfields

$$(2) \qquad \mathfrak{M} = \mathfrak{K}(x, y), \qquad \mathfrak{N} = \mathfrak{K}(z, x + yz), \qquad \mathfrak{R} = \mathfrak{K}(x)$$

are relatively algebraically closed in $\mathfrak{L}$, by Lüroth's theorem (cf. Steinitz [11], p. 126). The intersection of $\mathfrak{M}$ and $\mathfrak{N}$ is $\mathfrak{K}$. For let $\alpha \neq 0$ be an element of the intersection,

$$\alpha = f(x, y)/g(x, y) = r(t, z)/s(t, z), \qquad t = x + yz.$$

Then $g(x, y) \neq 0$, $s(t, z) \neq 0$, and we can assume that $r(t, z)$ and $s(t, z)$, as polynomials in t and z, have no factors in common except constants. Then $fs = gr$ is an identity in x, y and z; in it we set $z = 0$ to obtain

$$f(x, y)s(x, 0) = g(x, y)r(x, 0).$$

If $s(x, 0) = 0$, then, since $g(x, y) \neq 0$, $r(x, 0) = 0$. Since these are identities, $r(t, 0) = 0 = s(t, 0)$, which means that $r(t, z)$ and $s(t, z)$ have in common a factor z contrary to assumption. Hence $s(x, 0) \neq 0$ and

$$\alpha = f(x, y)/g(x, y) = r(x, 0)/s(x, 0),$$

so that α is in $\mathfrak{K}(x)$. If α were not in $\mathfrak{K}$, α would be transcendental over $\mathfrak{K}$ by Lüroth's theorem, so that x would be algebraic over $\mathfrak{K}(\alpha) \subset \mathfrak{N}$. Because $\mathfrak{N}$ is relatively algebraically closed, x is in $\mathfrak{N}$, and (2) shows that $\mathfrak{N}$ contains x, y, and z. This is a contradiction.

Because the intersection $\mathfrak{K} = \mathfrak{M} \cap \mathfrak{N}$ is the cross-cut $\mathfrak{M}\cdot\mathfrak{N}$ of $\mathfrak{M}$ and $\mathfrak{N}$ in the lattice L of relatively algebraically closed fields between $\mathfrak{K}$ and $\mathfrak{L}$, this lattice is not modular. For, by (2), $\mathfrak{M} > \mathfrak{R}$ and

$$\mathfrak{M}\cdot(\mathfrak{N} + \mathfrak{R}) = \mathfrak{M}\cdot\mathfrak{L} = \mathfrak{M}, \qquad \mathfrak{M}\cdot\mathfrak{N} + \mathfrak{R} = \mathfrak{K} + \mathfrak{R} = \mathfrak{R},$$

so that $\mathfrak{M}\cdot(\mathfrak{N} + \mathfrak{R}) \neq \mathfrak{M}\cdot\mathfrak{N} + \mathfrak{R}$, contrary to the Dedekind law.

This example can be extended in various ways. In any field $\mathfrak{K}(x_1, \cdots, x_n)$ with n independent variables one can construct two subfields $\mathfrak{M}$ and $\mathfrak{N}$, each of transcendence degree $n - 1$ over $\mathfrak{K}$, and with $\mathfrak{K}$ as intersection. In the field $\mathfrak{K}(x, y, z)$ one can also find a denumerable number of relatively algebraically closed subfields $\mathfrak{M}_i$, each of transcendence degree 2 over $\mathfrak{K}$, such that the intersection of any two of these subfields is $\mathfrak{K}$. These examples are typical of all fields, in the following sense:

THEOREM 11. *If $\mathfrak{K}$ is a relatively algebraically closed subfield of $\mathfrak{L}$, then the lattice of all relatively algebraically closed subfields $\mathfrak{M}$ with $\mathfrak{K} \subset \mathfrak{M} \subset \mathfrak{L}$ is a Dedekind lattice if and only if the transcendence degree of $\mathfrak{L}$ over $\mathfrak{K}$ is less than* 3.

Proof. When the transcendence degree is 1 or 2, the lattice has a simple form and the Dedekind law is trivially true. In the remaining cases, there are in $\mathfrak{L}$ three indeterminates x, y and z independent over $\mathfrak{K}$. Let $\mathfrak{M}$, $\mathfrak{N}$ and $\mathfrak{R}$ be the fields of (2) and $\mathfrak{M}'$, $\mathfrak{N}'$ and $\mathfrak{R}'$ be respectively their relative algebraic closures in $\mathfrak{L}$. Then $\mathfrak{M}' > \mathfrak{R}'$, and if we show that $\mathfrak{M}'$ and $\mathfrak{N}'$ intersect in $\mathfrak{K}$, the non-modularity follows as before.

Suppose then that w is an element common to $\mathfrak{M}'$ and $\mathfrak{N}'$, so that w is algebraic over both $\mathfrak{M}$ and $\mathfrak{N}$ and satisfies equations $f(u) = 0$ and $g(u) = 0$ irreducible over $\mathfrak{M}$ and $\mathfrak{N}$, respectively. Since $\mathfrak{K}(x, y, z)$ is a simple transcendental extension of $\mathfrak{M}$ and also of $\mathfrak{N}$, $f(u)$ and $g(u)$ must remain irreducible over the extension $\mathfrak{K}(x, y, z)$. If both f and g have the leading coefficient 1, then $f(u) \equiv g(u)$, so that w satisfies an equation $f(u) = 0$ whose coefficients are in $\mathfrak{M}\cdot\mathfrak{N}$; that is, in $\mathfrak{K}$. Therefore w, algebraic over $\mathfrak{K}$, is in $\mathfrak{K}$, and the intersection $\mathfrak{M}'\cdot\mathfrak{N}'$ is in fact $\mathfrak{K}$.

6. **Group ranks and p-bases.** Abelian groups furnish another example of exchange lattices. Let J be an additive Abelian group without elements of finite order (except 0). An element g of J is said to be dependent on a subset X of J if in X there are elements $x_1, \cdots, x_n$ and if there are integers $m, k_1, \cdots, k_n$, $m \neq 0$, with $mg = k_1x_1 + \cdots + k_nx_n$. This dependence relation has, as is readily verified, the five properties used in Theorem 1, so that the sets $H \subset J$ closed under this dependence form an exchange lattice. These closed subsets H are simply the subgroups $H \subset J$ for which the factor group J/H has no elements of finite order (except 0). A basis for J, in the lattice sense, is simply a

maximal independent set of elements of J, and the rank of J (Theorem 6) is a known invariant of J (Baer [1], Theorem 3.2). A simple proof shows also that the lattice of closed subgroups is a Dedekind lattice.

A field $\mathfrak{K}$ of characteristic p which is not perfect can be extended to a perfect field by the adjunction of p^n-th roots of a certain minimal set of elements of $\mathfrak{K}$. This minimal set is called a p-basis (Teichmüller [12], §3, and [13], p. 145, Hilfsatz 9). This notion can be generalized to any inseparable extension.[6] Let $\mathfrak{K}$ be a field of characteristic p, and $\mathfrak{L}$ a pure inseparable extension of exponent 1 over $\mathfrak{K}$; that is, an extension such that x^p is in $\mathfrak{K}$ for any x in $\mathfrak{L}$. An element y in $\mathfrak{L}$ will be called *p-dependent* on a subset X of $\mathfrak{L}$ if y is in the field $\mathfrak{K}(X)$. The properties 1, 2, 3 and 5 of a dependence relation are immediately verified. As for property 4, let y be in $\mathfrak{K}(x_1, \cdots, x_n)$, but not in $\mathfrak{K}(x_1, \cdots, x_{n-1})$. Since x_n^p is in $\mathfrak{K}$, $y = f(x_1, \cdots, x_n)$, where f is a polynomial with coefficients in $\mathfrak{K}$ and of degree less than p in x_n. Then x_n must actually occur in some term of f, so x_n is algebraic of degree $\leqq p - 1$ over $\mathfrak{K}_1 = \mathfrak{K}(y, x_1, \cdots, x_{n-1})$. But x_n also satisfies the inseparable equation $z^p - x_n^p = 0$ over $\mathfrak{K}_1$, so that x_n must be in $\mathfrak{K}_1$, and is p-dependent on $y, x_1, \cdots, x_{n-1}$, as asserted in property 4.

The exchange lattice corresponding to this dependence relation is simply the lattice of all subfields $\mathfrak{M}$ with $\mathfrak{K} \subset \mathfrak{M} \subset \mathfrak{L}$. The points of the lattice are the subfields of degree p over $\mathfrak{K}$, so that Theorem 6 becomes

THEOREM 12. *If $\mathfrak{L}$ is a pure inseparable extension of exponent 1 of a field $\mathfrak{K}$ of characteristic p, then there exists a set of subfields $\mathfrak{M}_\sigma$ of $\mathfrak{L}$, each of degree p over $\mathfrak{K}$, such that the adjunction of all $\mathfrak{M}_\sigma$ to $\mathfrak{K}$ gives $\mathfrak{L}$, while no $\mathfrak{M}_\sigma$ is a subfield of the field obtained by adjoining the remaining $\mathfrak{M}$'s to $\mathfrak{K}$. The cardinal number of subfields $\mathfrak{M}_\sigma$ is an invariant of $\mathfrak{L}/\mathfrak{K}$, called the relative degree of imperfection.*

Each subfield $\mathfrak{M}_\sigma$ has the form $\mathfrak{M}_\sigma = \mathfrak{K}(\sqrt[p]{x_\sigma})$ for some x_σ in $\mathfrak{K}$. The set of these x_σ's, one for each $\mathfrak{M}_\sigma$, can be called a *p-basis* of $\mathfrak{L}$ over $\mathfrak{K}$. This concept will apply to any inseparable algebraic extension, for such an extension $\mathfrak{F}$ of $\mathfrak{K}$ can be uniquely decomposed into $\mathfrak{K} \subset \mathfrak{L}_0 \subset \mathfrak{L}_1 \subset \mathfrak{L}_2 \subset \cdots \subset \mathfrak{F}$, where each $\mathfrak{L}_n$ contains all elements of $\mathfrak{F}$ of exponent not more than n over $\mathfrak{K}$, so that each $\mathfrak{L}_n$ is a pure inseparable extension of exponent 1 over $\mathfrak{L}_{n-1}$.

The lattice of subfields $\mathfrak{M}$ of $\mathfrak{L}$ is not always modular. Consider $\mathfrak{L} = \mathfrak{P}(x, y, z)$, $\mathfrak{K} = \mathfrak{P}(x^p, y^p, z^p)$, where x, y and z are independent indeterminates over the perfect field $\mathfrak{P}$ of characteristic p. $\mathfrak{L}$ over $\mathfrak{K}$ is pure inseparable of degree p^3, is obtained from $\mathfrak{K}$ by adjoining the independent p-th roots x, y and z, and so has the p-basis $\{x, y, z\}$ (cf. Teichmüller [12], Theorem 18). The subfields $\mathfrak{M} = \mathfrak{K}(x, y)$, $\mathfrak{N} = \mathfrak{K}(z, x + yz)$ are each of degree p^2 over $\mathfrak{K}$. To disprove the Dedekind law, it will suffice as in §5 to show that $\mathfrak{M}$ and $\mathfrak{N}$ intersect in $\mathfrak{K}$. Let them have in common the element

$$\alpha = f(x, y) = g(z, t) \equiv \sum_{i,j=0}^{p-1} a_{ij} z^i t^j, \qquad t = x + yz, \tag{3}$$

[6] The possibility of such an extension is indicated by Teichmüller [12], §3.

where f and g are polynomials with coefficients in $\mathfrak{K}$ and of degree less than p in any one variable. The polynomial

$$h(x, y, z) \equiv f(x, y) - g(z, x + yz) \tag{4}$$

must be zero in $\mathfrak{L}$ and is of degree at most $2p - 2$ in the variable z. If each term z^{p+s} be replaced by dz^s, with coefficient $d = z^p$ in $\mathfrak{K}$, then $h(x, y, z)$ is equal in $\mathfrak{L}$ to a new polynomial $h'(x, y, z)$ of degree less than p in any one variable. But the power products $x^q y^r z^s$ with exponents less than p form a basis for the algebraic extension $\mathfrak{L}/\mathfrak{K}$, so that $h'(x, y, z) = 0$ in $\mathfrak{L}$ implies that h' is identically zero. In h' terms in z^{p-1} never arise from a replacement $z^{p+s} \rightarrow dz^s$, but come only from terms $z^i t^j$ in g with $i + j \geqq p - 1$. These terms are, by expansion of (3),

$$\left(\sum_{j=0}^{p-1} \sum_{i=p-1-j}^{p-1} \binom{j}{p-1-i} a_{ij} x^{j+i-(p-1)} y^{p-1-i}\right) z^{p-1}.$$

These terms involve distinct power products $x^q y^s$ and have binomial coefficients not zero, so $h' \equiv 0$ implies $a_{ij} = 0$ for $i + j \geqq p - 1$. Thus (4) actually involves no term of degree p or more in x, in y or in z, and $h(x, y, z) \equiv 0$. But z arises only from $g(z, x + yz)$, so only the constant term of g can differ from 0, and the element $\alpha = g(0, 0)$ of $\mathfrak{M} \cdot \mathfrak{N}$ is in fact in $\mathfrak{K}$.

7. **Exchange axioms free of points**. A central feature of von Neumann's continuous geometry is the use of the modular law, which makes no reference to the points of the geometry. Similarly, Wilcox [16] has shown that affine geometry as developed algebraically by Menger can also be axiomatized without the use of points. His treatment depends on certain properties of a relation of modularity which do not hold in all exchange lattices. Nevertheless, our exchange axiom can be replaced by conditions which make no use of points or of covering relations.

To modify Menger's exchange axiom (E_3), which asserts that $p \nleqq a + c$ implies $(a + p)c = ac$, replace the point p by an arbitrary element b and the conclusion $(a + p)c = ac$ by the assertion that $(a + b_1)c = ac$ for some non-trivial part b_1 of b. No generality is lost if we require $a < c$; the hypothesis $p \nleqq a + c$ or $p(a + c) = 0$ might then become $bc < a$, and our modified statement is[7]

(E_5) $bc < a < c < b + c$ implies that there exists b_1 such that $bc < b_1 \leqq b$ and $(a + b_1)c = a$.

A similar elimination of points from the exchange axiom (E_1) leads eventually to the laws

(E_6) $bc < a < c < b + c$ implies that there exists b_1 such that $bc < b_1 \leqq b$ and $(a + b_1)c < c$;

(E_7) $bc < a < c < b + a$ implies the conclusion of (E_6).

The second assertion $(a + b_1)c < c$ of this conclusion is equivalent to $c \nleqq a + b_1$, which in turn is equivalent to $a + b_1 < c + b_1$.

[7] The statement that this law (or other similar laws) holds is to mean that it holds for all elements a, b, and c of the lattice.

These three point-free laws do not, like the exchange axiom (E_1), hold trivially in any lattice without points (as for instance in the lattice of all real numbers between 0 and 1, where $a < b$ has its usual meaning). The rôle of these laws can be stated thus:

THEOREM 13. *In any lattice,* (E_5) *implies* (E_6) *and* (E_6) *is equivalent to* (E_7).

THEOREM 14. *In a lattice satisfying the point-existence axiom* (G_1), *the exchange axiom* (E_1) *is equivalent to* (E_5) *and also to* (E_6). *Hence* (E_5) (*or* (E_6)) *can replace* (E_1) *in the definition of an exchange lattice.*

THEOREM 15. *In any lattice,* (E_7) (*and hence* (E_6) *or* (E_5)) *implies the covering law* (E_4). *Therefore a dimension function* $\rho(a)$ *can be defined in any complete lattice of finite dimensions satisfying* (E_7); *in other words, the conclusions of Theorem 9 hold for such a lattice.*

Proof. $(E_5) \rightarrow (E_6) \rightarrow (E_7)$ is immediate. Conversely, to prove $(E_7) \rightarrow (E_6)$, let $bc < a < c < b + c$ as in (E_6). If $a + b > c$, then the hypothesis of (E_7) holds and yields the desired conclusion. If $a + b = c$, then $b \leqq c$, $b + c = c < b + c$, a contradiction. In the remaining case, $(a + b)c < c$, which states that the conclusion of (E_6) holds with $b_1 = b$. This gives Theorem 13.

For Theorem 15 we need only prove $(E_7) \rightarrow (E_4)$. Given the hypothesis of (E_4), the conclusion could be false only if $a + b > c > a$ for some c. Then $d \leqq bc \leqq b$, whence $bc = b$ or $bc = d$. In the former case, $b \leqq c$, $a + b \leqq a + c = c < a + b$, a contradiction. Therefore $bc = d$. Omitting the trivial case $d = a$, we have $bc < a < c < a + b$, as in (E_7), so that $(a + b_1)c < c$ for $bc < b_1 \leqq b$. Because b covers bc, $b_1 = b$ and $c = (a + b)c = (a + b_1)c < c$, a contradiction.

Since $(E_7) \rightarrow (E_4)$, Theorem 14 now needs only a proof that (E_5) holds whenever (E_1) and (G_1) do. Let $bc < a < c < b + c$. Since $b \nleqq c$, the property (G_2) of §2 furnishes a point p with $p \leqq b$, $p \nleqq c$. Hence $p \nleqq a + c = c$, so that $(a + p)c = ac$ by the exchange axiom (E_3), which is known to hold (Theorem 8). If we set $b_1 = bc + p$, then $bc < b_1 \leqq b$, while $(a + b_1)c = (a + bc + p)c = (a + p)c = ac = a$, as in the conclusion of (E_5).

Any one of these three point-free axioms can be viewed as a weaker form of the modular law (D_1), for in any lattice this modular law is equivalent to the following assertion, of a form similar to (E_6),

(D_2) $bc < a < c < b + c$ implies $(a + b)c = a$.

This is a direct consequence of the modular law. Conversely, the modular law requires that $c' = c(b + a)$ be equal to $a' = cb + a$ for any $c \geqq a$. In any event, $c' \geqq a'$ and $bc' \leqq a'$. If either $bc' = a'$ or $b + c' = c'$, the conclusion $c' = a'$ will follow readily. Hence the conclusion can be false only with $bc' < a' < c' < b + c'$, exactly as in the hypothesis of (D_2), whence we obtain the conclusion

$$cb + a = a' = (a' + b)c' = (cb + a + b)c(b + a)$$
$$= (a + b)c(b + a) = c(b + a),$$

which is the modular law.

The complexity of the axioms (E_5) to (E_7), attendant on the existence assertion for b_1, is unavoidable. Specifically, suppose that an axiom (E_M) could be found which, like (E_5), is equivalent to the exchange axiom (E_1) in the presence of the axioms (F_1) and (G_1), but which, unlike (E_5), contains no existence assertions. In other words, (E_M) is built up from certain elements $a_1, \cdots, a_n$ of the lattice L by any combinations using $<$, $=$, $+$, $\cdot$, "and", "implies" and "not", and is to be asserted for all choices of the elements a_i. Any such axiom (E_M) true in a lattice L is automatically true in any sublattice L' of L, for $<$, $+$, and $\cdot$ have the same meaning in L' as in L. But then (E_M) cannot be equivalent to (E_1). For the non-modular lattice L_{42} of §4 contains four points p, q, r and s and has a sublattice L' composed of 0, p, r, $p + r$, $q + s$ and 1. In L' axioms (F_1) and (G_1) hold, but axiom (E_1) fails, because in L' p and $q' = q + s$ are points, $p \leqq r + q'$, $p \nleqq r$, so that, according to (E_1), $q' = q + s \leqq r + p$ should hold, contrary to the construction of L_{mn}. But (E_M), true in L, is also true in L', and so is not equivalent to (E_1) in the presence of (F_1) and (G_1).

8. **Transposition axioms and modularity**. The point-free exchange axiom (E_5) states in part that the chain in its hypothesis can be subjected to the following successive transpositions,

$$bc < a < c < b + c, \tag{5}$$

$$bc < a < a + b_1 < b + c, \tag{6}$$

$$bc < b_1 < a + b_1 < b + c. \tag{7}$$

Note that the inclusions in (6) are all proper inclusions, for $a + b_1 = a$ would give $b_1 \leqq a$, $bc < b_1 \leqq ab \leqq cb$, a contradiction to ($E_5$), while $a + b_1 = b + c$ implies $(a + b_1)c = (b + c)c = c$, although $(a + b_1)c = a$ in (E_5). Similarly $b_1 < a + b_1$ in (7).

To investigate the relations of the exchange axiom to the possibility of such transpositions, consider any chain of length n,

$$C : a_0 < a_1 < a_2 < \cdots < a_n . \tag{8}$$

A *transposition* of C is the operation of replacing an element a_k by a new element a_k' between a_{k-1} and a_{k+1} ($a_{k-1} < a_k' < a_{k+1}$). We call the transposition *primary* if $a_k \cdot a_k' = a_{k-1}$, *proper* if $a_k \cdot a_k' < a_k$ and $a_k \cdot a_k' < a_k'$. A primary transposition is always proper, but not conversely. In these terms can be stated the following alternative modular and exchange axioms:

(T_5) If C is a chain of length $n \geqq 2$ and if $a_0 < b < a_n$ and $a_1 b = a_0$, then[8] one can find $n - 1$ or fewer successive primary transpositions of C which yield a chain $a_0 < b_1 < \cdots < a_n$ with second term $b_1 \leqq b$.

[8] The hypothesis $a_1 b = a_0$ is not highly restrictive, since in the contrary case $a_1 \geqq a_1 b > a_0$, so a simple insertion of $b_1 = a_1 b$ in C yields a new chain of the desired form. If in (T_5) the hypothesis $a_1 b = a_0$ were replaced by $a_{n-1} b = a_0$, and the conclusion "$n - 1$ or fewer" by "exactly $n - 1$", the resulting statement is equivalent to the original (T_5), by the arguments used for Theorem 16.

(T_6) Statement as in (T_5), with "primary" replaced by "proper".

(T_D) If C is a chain of length $n \geqq 2$ and if $a_0 < b < a_n$ and $a_{n-1}b = a_0$, then one can find $n - 1$ successive primary transpositions of C which yield a chain $a_0 < b < \cdots < a_n$ with second term b.

THEOREM 16. *In any lattice L, the Dedekind law is equivalent to* (T_D), *while the exchange axioms* (E_5) *and* (E_6) *are equivalent respectively to* (T_5) *and* (T_6).

We prove (E_5) $\to$ (T_5) by induction. If $n = 2$, C is $a_0 < a_1 < a_2$, and the transposition to $a_0 < b < a_2$ is primary by the hypothesis $a_1b = a_0$. Let (T_5) be true for all chains in L of length $k < n$. If $a_{n-1}b > a_0$ in the given chain C, then $a_{n-1} > a_{n-1}b > a_0$, so we apply the induction assumption to $b' = a_{n-1}b$ and the chain $a_0 < \cdots < a_{n-1}$ of length $n - 1$. In the remaining case, $a_{n-1}b = a_0$, $b \nleqq a_{n-1}$, so the four-element chain $a_0 < a_{n-2} < a_{n-1} < a_{n-1} + b$ has the form of the hypothesis of (E_5). There exists a b_1, $a_0 < b_1 \leqq b$, with $(a_{n-2} + b_1)a_{n-1} = a_{n-2}$, which is to say that the transposition $a_{n-1} \to (a_{n-2} + b_1)$ in C is primary. The induction assumption applied to the chain $a_0 < a_1 < \cdots < a_{n-2} < a_{n-2} + b_1$ and to the element b_1 therefore gives $n - 2$ more transpositions leading to the desired type of chain.

As to the converse, (T_5) $\to$ (E_5), the chain (5) in the hypothesis of (E_5) is by (T_5) reducible by one or two primary transpositions to the form $bc < b_1 < c_1 < b + c$, with $b_1 \leqq b$. The first of these transpositions could not have introduced b_1, for then $b_1 < c$, which entails the contradiction $b_1 = bb_1 \leqq bc < b_1$. The first transposition is therefore $c \to c_1$; it is primary, so $cc_1 = a$. But $a < c_1$, $b_1 < c_1$, $a + b_1 \leqq c_1$ so $c(a + b_1) \leqq a$, and therefore $c(a + b_1) = a$, which is the conclusion of (E_5). This argument depends essentially on the circumstance that (T_5) allows only *two* transpositions on the chain (5).

An analogous proof that (T_6) $\leftrightarrow$ (E_6) is possible because (E_6) again gives two transpositions (5) $\to$ (6) and (6) $\to$ (7) in the chain (5) of length 3. The second transposition is primary, hence proper. The first transposition is proper because (E_6) asserts that $(a + b_1)c < c$ and because $(a + b_1)c < a + b_1$, for otherwise $a + b_1 \leqq c$ would give the paradox $bc \geqq b_1 > bc$.

To show that the Dedekind law implies (T_D), observe that the chain (8) can be subjected to the successive transpositions $a_k \to (a_{k-1} + b)$ $(k = n - 1, \cdots, 1)$. Each one is primary, for by the Dedekind law, $(a_{k-1} + b)a_k = a_{k-1} + ba_k \leqq a_{k-1} + ba_{n-1} = a_{k-1} + a_0 = a_{k-1}$. The final chain $a_0 < b < a_1 + b < \cdots < a_n$ has the specified form. The converse assertion that (T_D) $\to$ (D_1) can be readily checked as above by using the Dedekind law in the form (D_2) of §7.

BIBLIOGRAPHY

1. R. BAER, *The subgroup of elements of finite order of an Abelian group*, Annals of Mathematics, vol. 37(1936), p. 768.
2. G. BIRKHOFF, *On the combination of subalgebras*, Proceedings of the Cambridge Philosophical Society, vol. 29(1933), pp. 441–464.
3. G. BIRKHOFF, *Abstract linear dependence and lattices*, American Journal of Mathematics, vol. 57(1935), pp. 800–804.
4. G. BIRKHOFF, *Combinatorial relations in projective geometries*, Annals of Mathematics, (2), vol. 36(1935), pp. 743–748.

5. G. Köthe, *Die Theorie der Verbände*, Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 47(1937), pp. 125–144.
6. S. Mac Lane, *Some interpretations of abstract linear dependence in terms of projective geometry*, American Journal of Mathematics, vol. 58(1936), pp. 236–240.
7. K. Menger, *New foundations of projective and affine geometry; algebra of geometry*, Annals of Mathematics, (2), vol. 37(1936), pp. 456–482.
8. J. von Neumann, *Continuous Geometry*, Notes of lectures at the Institute for Advanced Study, 1936.
9. O. Ore, *On the foundation of abstract algebra*. I, Annals of Mathematics, (2), vol. 36 (1935), pp. 406–437.
10. O. Ore, *On the decomposition theorems of algebra*, Comptes Rendus du Congrès International des Mathématiciens, 1936.
11. E. Steinitz, *Algebraische Theorie der Körper*, edited by R. Baer and H. Hasse, 1930.
12. O. Teichmüller, *p-Algebren*, Deutsche Mathematik, vol. 1(1936), pp. 362–388.
13. O. Teichmüller, *Diskret bewertete perfekte Körper mit unvollkommenem Restklassenkörper*, Journal für die Mathematik, vol. 176(1936), pp. 141–152.
14. B. L. van der Waerden, *Moderne Algebra*, vol. I, p. 204.
15. H. Whitney, *On the abstract properties of linear dependence*, American Journal of Mathematics, vol. 57(1935), pp. 509–533.
16. L. R. Wilcox, *Modularity in the theory of lattices*, Bulletin of the American Mathematical Society, vol. 44(1938), p. 50 (abstract 44-1-79).

University of Chicago.

PROOF OF A GAP THEOREM

By J. Marcinkiewicz and A. Zygmund

Using the theory of Fourier transforms, Wiener[1] proved the following

Theorem. *Let us suppose that the trigonometrical series*

$$\frac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k \cos \lambda_k x + b_k \sin \lambda_k x), \tag{1}$$

where $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$, *is quadratically bounded over an interval* (a, b), *that is, that there exists a number* M *such that*

$$\int_a^b s_n^2(x)\, dx \leqq M^2, \tag{2}$$

where

$$s_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos \lambda_k x + b_k \sin \lambda_k x)$$

for $n = 1, 2, \cdots$.[2] *Let*

$$\lambda_n - \lambda_{n-1} \geqq \Delta > 0 \qquad (n = 1, 2, \cdots). \tag{3}$$

We write $b - a = \delta$. *Then, if* Δ *is sufficiently large,*

$$\Delta \geqq \Delta_0 = \Delta_0(\delta), \tag{4}$$

the series $\sum (a_k^2 + b_k^2)$ *converges, and*

$$\frac{1}{2}a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leqq A(\delta)M^2, \tag{5}$$

where $A(\delta)$ *is a constant depending only on* δ.[3]

The object of this note is to give a new and elementary proof of this theorem.

Received March 1, 1938.

[1] N. Wiener, *A class of gap theorems*, Annali di Pisa, vol. 2(1934), pp. 367–372.

[2] We may restrict ourselves to the case of real coefficients.

[3] The numbers λ_i need not be integers. If $\lambda_1, \lambda_2, \cdots$ are integers, the theorem may also be stated as follows:

Under the conditions (2), (3), (4), *the series* (1) *converges in mean to a function* $f(x)$ *such that*

$$\frac{1}{\pi}\int_0^{2\pi} |f(x)|^2\, dx \leqq A(\delta) \int_a^b |f(x)|^2\, dx.$$

Without loss of generality we may assume that (a, b) is of the form $(-\frac{1}{2}\delta, \frac{1}{2}\delta)$. We write the polynomial $s_n(x)$ in the form

$$s_n(x) = \frac{1}{2} \sum_{k=-n}^{n} c_k e^{i\lambda_k x},$$

where

$$c_k = a_k - ib_k, \qquad c_{-k} = \bar{c}_k, \qquad \lambda_{-k} = -\lambda_k$$

for $k > 0$, and consider the integral

$$\begin{aligned} \int_{-h}^{h} s_n^2(x) &= \int_{-h}^{h} \left\{\frac{1}{2} \sum_{k=-n}^{n} c_k e^{i\lambda_k x}\right\}\left\{\frac{1}{2} \sum_{l=-n}^{n} \bar{c}_l e^{-i\lambda_l x}\right\} dx \\ &= \frac{1}{2} h \sum_{k=-n}^{n} | c_k |^2 + \frac{1}{4} \int_{-h}^{h} \left\{ \sum_{\substack{k,l=-n \\ k \neq l}}^{n} c_k \bar{c}_l e^{i(\lambda_k - \lambda_l)x} \right\} dx. \end{aligned} \tag{6}$$

Here h is any positive number not exceeding $\frac{1}{2}\delta$. The last integral on the right side of (6) we denote by $I(h)$. Hence

$$I(h) = \frac{1}{2} \sum_{k,l=-n}^{n}{}' c_k \bar{c}_l \frac{\sin (\lambda_k - \lambda_l)h}{\lambda_k - \lambda_l},$$

$$\begin{aligned} \left| \int_{\frac{1}{4}\delta}^{\frac{1}{2}\delta} I(h)\, dh \right| &\leqq \sum_{k,l=-n}^{n}{}' \frac{| c_k c_l |}{(\lambda_k - \lambda_l)^2} \\ &\leqq \frac{1}{2} \sum_{k,l=-n}^{n}{}' \frac{| c_k |^2 + | c_l |^2}{(\lambda_k - \lambda_l)^2} = \sum_{k=-n}^{n} | c_k |^2 \sum_{l=-n}^{n}{}' \frac{1}{(\lambda_k - \lambda_l)^2}, \end{aligned} \tag{7}$$

where the $\sum'$ indicates that the terms with $l = k$ are omitted. From (3) we see that the coefficient of $| c_k |^2$ does not exceed $\frac{1}{3}\pi^2\Delta^{-2}$. This and formula (7) give

$$\left| \int_{\frac{1}{4}\delta}^{\frac{1}{2}\delta} I(h)\, dh \right| \leqq \frac{\pi^2}{3\Delta^2} \sum_{k=-n}^{n} | c_k |^2. \tag{8}$$

If $I(h)$ is of constant sign in the interval $(\frac{1}{4}\delta, \frac{1}{2}\delta)$, the inequality (8) shows that this interval contains a number ξ such that

$$\frac{1}{4}\delta \, | I(\xi) | \leqq \frac{\pi^2}{3\Delta^2} \sum_{k=-n}^{n} | c_k |^2. \tag{9}$$

If $I(h)$ is not of constant sign in the interval $(\frac{1}{4}\delta, \frac{1}{2}\delta)$, then $I(h)$ vanishes at a point $h = \xi$ of the interval, so that we have (9) again. Make $h = \xi$ in (6). On account of (9) we obtain

$$\begin{aligned} \int_{-\frac{1}{2}\delta}^{\frac{1}{2}\delta} s_n^2(x)\, dx \geqq \int_{-\xi}^{\xi} s_n^2(x)\, dx &\geqq \left(\frac{1}{2}\xi - \frac{4\pi^2}{3\Delta^2\delta}\right) \sum_{k=-n}^{n} | c_k |^2 \\ &\geqq \left(\frac{1}{8}\delta - \frac{4\pi^2}{3\Delta^2\delta}\right) \sum_{k=-n}^{n} | c_k |^2. \end{aligned} \tag{10}$$

If we assume that

$$\frac{4\pi^2}{3\Delta^2\delta} \leqq \frac{2}{3}\cdot\frac{1}{8}\delta, \qquad \text{that is, } \Delta \geqq 2\cdot\frac{2\pi}{\delta},$$

we see from (10) that

$$\sum_{k=-n}^{n} |c_k|^2 \leqq \frac{24}{\delta}\int_{-\frac{1}{2}\delta}^{\frac{1}{2}\delta} s_n^2(x)\,dx \leqq \frac{24}{\delta}M^2. \tag{11}$$

On allowing n to become infinite, we obtain the inequality (5) with $A(\delta) = 12/\delta$ and $\Delta_0 = 4\pi/\delta$. This completes the proof of the theorem.

Remarks. (i) It is plain that for the proof of the convergence of the series $\sum (a_k^2 + b_k^2)$ it is sufficient to suppose only that the condition (3) is satisfied for all sufficiently large values of n. For then, omitting a finite number of terms of the series (1) (which may only influence the value of the constant M), we obtain a series for which the condition (3) is satisfied for all n.

(ii) Wiener proved his result in a slightly more general form, viz., *for the validity of* (5) *it is sufficient to suppose instead of* (2) *that the Abel means of the series* (1) *are quadratically bounded over the interval* (a, b), *that is, that*

$$\int_a^b f^2(\rho, x)\,dx \leqq M^2 \quad \text{for } 0 \leqq \rho < 1,$$

where

$$f(\rho, x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty}(a_k \cos \lambda_k x + b_k \sin \lambda_k x)\rho^{\lambda_k}.$$

For the proof of this theorem it is sufficient to observe that the series defining $f(\rho, x)$ is absolutely convergent for $0 \leqq \rho < 1$, and so, arguing as above, we obtain instead of (11) the inequality

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \rho^{2|\lambda_k|} \leqq \frac{24}{\delta}M^2.$$

Allowing here ρ to tend to 1, we obtain (5).

(iii) In his paper cited, Wiener gave a number of applications of the theorem we have established. We shall give here one more application, viz.

The entire function

$$f(z) = \sum_{k=0}^{\infty} c_k z^{\lambda_k},$$

where $\lambda_k - \lambda_{k-1} \to \infty$, *is of the same order and of the same type in every angle* $a \leqq \arg z \leqq b$.

For let us suppose that

$$|f(re^{i\theta})| = O(e^{\beta r^\alpha}) \qquad (a \leqq \theta \leqq b).$$

This inequality holds if we replace $|f(re^{i\theta})|$ on the left side by $\left(\int_a^b |f(re^{i\theta})|^2 d\theta\right)^{\frac{1}{2}}$, and so, in view of Wiener's theorem,

$$\left(\int_0^{2\pi} |f(re^{i\theta})|^2 d\theta\right)^{\frac{1}{2}} = O(e^{\beta r^\alpha}).$$

Let $F(z) = U(z) + iV(z)$ be the integral of $f(z)$ vanishing at the origin. Since the integrals of the functions U and V over the interval $0 \leqq \theta \leqq 2\pi$ are equal to 0, we have $U(re^{i\theta_1}) = V(re^{i\theta_2}) = 0$ for some θ_1 and θ_2, and so

$$\begin{aligned} |F(re^{i\theta})| &\leqq \left|\int_{\theta_1}^{\theta} \frac{d}{d\theta} U(re^{i\theta})\, d\theta\right| + \left|\int_{\theta_2}^{\theta} \frac{d}{d\theta} V(re^{i\theta})\, d\theta\right| \\ &\leqq 2r \int_0^{2\pi} |f(re^{i\theta})|\, d\theta \leqq 2\sqrt{2\pi} r \left(\int_0^{2\pi} |f(re^{i\theta})|^2 d\theta\right)^{\frac{1}{2}} = O(re^{\beta r^\alpha}) = O(e^{\beta_1 r^\alpha}) \end{aligned}$$

for $0 \leqq \theta \leqq 2\pi$ and $\beta_1 > \beta$. On comparing the extreme terms of these inequalities and making use of the well known fact that neither the order nor the type of an entire function is changed by differentiation, we obtain the required result.

UNIVERSITY OF WILNO.

A THEOREM OF LUSIN

BY J. MARCINKIEWICZ AND A. ZYGMUND

Part I

1. Let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{1}$$

be a function holomorphic in the circle $|z| < 1$. The function $f(z)$ is said to belong to the class H^λ, $\lambda > 0$, if the expression

$$I_\lambda(r) = I_\lambda(r, f) = \frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^\lambda \, d\theta \tag{2}$$

is bounded for $r < 1$. It is well known[1] that, if $f(z)$ belongs to H^λ, then for almost every θ the limit

$$f(e^{i\theta}) = \lim_{z \to e^{i\theta}} f(z) \tag{3}$$

exists, where z tends to $e^{i\theta}$ along any non-tangential path. Hence, if C denotes the upper bound of the expression (2) for $0 \leqq r < 1$, we have

$$\int_0^{2\pi} |f(e^{i\theta})|^\lambda \, d\theta \leqq C. \tag{4}$$

In the sequel we shall also use the fact that the expression $I_\lambda(r)$ is a non-decreasing function of r, and so in particular

$$I_\lambda(r) \leqq \frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^\lambda \, d\theta \qquad (0 \leqq r < 1). \tag{5}$$

If $\lambda \geqq 1$, the real part and the imaginary part of the power series (1) on the circle $|z| = 1$ are both Fourier series of functions of the class L^λ.

Let Ω denote the interior of a simple closed curve Γ given by the equation

$$\rho = \psi(\theta) \qquad (-\pi \leqq \theta \leqq \pi) \tag{6}$$

and possessing, among others, the following two properties:

(i) Γ passes through the point $z = 1$, but otherwise lies entirely in the circle $|z| < 1$;

(ii) Γ is not tangent to the circle $|z| = 1$ at the point $z = 1$, that is, there

Received March 1, 1938.

[1] See F. Riesz, *Über die Randwerte einer analytischen Funktion*, Math. Zeitschr., vol. 18 (1922), pp. 87–95.

exist two positive numbers ϵ and δ such that, for every point z belonging to Γ and satisfying the inequality $|z| \geqq 1 - \epsilon$, we have

$$|\arg(1-z)| \leqq \tfrac{1}{2}\pi - \delta.$$

We shall also suppose for simplicity (which will in no way cause any restriction of generality of our considerations) that

(iii) Γ is symmetric with respect to the real axis,

(iv) $\psi(\theta)$ is a decreasing function of θ in the interval $0 \leqq \theta \leqq \pi$.

Let Γ_u denote the curve Γ rotated about the point $z = 0$ through an angle u. Hence Γ_u has only the point $z = e^{iu}$ in common with circle $|z| = 1$. We shall denote the interior of the curve Γ_u by Ω_u.

The following theorem has been established by Lusin.[2]

THEOREM A. *If $f(z)$ belongs to H^2, then for almost every u the integral*

$$S(u) = S(u,f) = \iint_{\Omega_u} |f'(z)|^2\, dx\, dy \qquad (z = x + iy) \tag{7}$$

is finite. Moreover, $S(u)$ is an integrable function of u satisfying an inequality

$$\int_0^{2\pi} S(u)\, du \leqq A \int_0^{2\pi} |f(e^{i\theta})|^2\, d\theta,$$

where the constant A depends only on the curve Γ.

Besides $S(u)$ we shall also consider the function

$$s(u) = s(u,f) = S^{\frac{1}{2}}(u,f) = \left(\iint_{\Omega_u} |f'(z)|^2\, dx\, dy\right)^{\frac{1}{2}}.$$

The main purpose of Part I of this paper is to establish the following proposition.

THEOREM 1. *If $f(z)$ belongs to H^λ, $\lambda > 0$, the function $s(u)$ is finite for almost every u and belongs to L^λ. More precisely,*

$$\left(\int_0^{2\pi} s^\lambda(u)\, du\right)^{1/\lambda} \leqq A_{\lambda,\Gamma}\left(\int_0^{2\pi} |f(e^{i\theta})|^\lambda\, d\theta\right)^{1/\lambda}, \tag{8}$$

where $A_{\lambda,\Gamma}$ depends only on λ and Γ.

2. The argument which follows is based mainly on certain results of Hardy, Littlewood and Paley.

LEMMA 1. *Let $f(z)$ belong to the class H^λ, $\lambda > 1$, in the circle $|z| < 1$, and let*

$$f(0) = 0. \tag{9}$$

We write

$$g(\theta) = \left(\int_0^1 (1-\rho)\,|f'(\rho e^{i\theta})|^2\, d\rho\right)^{\frac{1}{2}}. \tag{10}$$

[2] N. Lusin, *Sur une propriété des fonctions à carré sommable*, Bulletin of the Calcutta Mathematical Society, vol. 20(1930), pp. 139–154.

Then the function $g(\theta)$ belongs to the class L^λ and satisfies an inequality

$$\left(\int_0^{2\pi} g^\lambda(\theta)\,d\theta\right)^{1/\lambda} \leqq B_\lambda \left(\int_0^{2\pi} |f(e^{i\theta})|^\lambda\,d\theta\right)^{1/\lambda},$$

where B_λ depends on λ only.

This result is due to Littlewood and Paley.[3]

LEMMA 2. *Let $f(z)$ belong to H^λ, $\lambda > 0$, and let*

$$M(u) = M(u, f) = \max |f(z)|$$

for z belonging to Ω_u. Then

$$\left(\int_0^{2\pi} M^\lambda(u)\,du\right)^{1/\lambda} \leqq C_{\lambda,\Gamma} \left(\int_0^{2\pi} |f(e^{i\theta})|^\lambda\,d\theta\right)^{1/\lambda},$$

where $C_{\lambda,\Gamma}$ depends only on λ and Γ.

LEMMA 3. *Let $h(\theta)$ be a function of period 2π, belonging to the class L^p over the interval $(0, 2\pi)$, where $p > 1$. Let*

$$h^*(\theta) = \max_{\theta_1 \leqq \theta \leqq \theta_2} \left| \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} h(u)\,du \right|.$$

Then

$$\int_0^{2\pi} \{h^*(\theta)\}^p\,d\theta \leqq 2\left(\frac{p}{p-1}\right)^p \int_0^{2\pi} |h(\theta)|^p\,d\theta. \tag{11}$$

Lemmas 2 and 3 are also known.[4]

3. We now pass on to the proof of Theorem 1. We begin with the case $\lambda > 2$ and suppose first that the condition (9) is satisfied. Let μ denote the exponent conjugate to $\frac{1}{2}\lambda$, that is,

$$\frac{1}{\frac{1}{2}\lambda} + \frac{1}{\mu} = 1. \tag{12}$$

We then may write

$$\left(\int_0^{2\pi} s^\lambda(u)\,du\right)^{1/\lambda} = \left(\int_0^{2\pi} S^{\frac{1}{2}\lambda}(u)\,du\right)^{1/\lambda} = \max_{h \in X_\mu} \left(\int_0^{2\pi} S(u)h(u)\,du\right)^{\frac{1}{2}}, \tag{13}$$

where X_μ denotes the class of functions $h(u)$ such that

$$\int_0^{2\pi} |h(u)|^\mu\,du \leqq 1. \tag{14}$$

[3] J. E. Littlewood and R. E. A. C. Paley, *Theorems on Fourier series and power series, II*, Proc. London Math. Soc., vol. 42(1937), pp. 52–89.

[4] G. H. Hardy and J. E. Littlewood, *A maximal theorem with function-theoretic applications*, Acta Mathematica, vol. 54(1930), pp. 81–116.

We take for h any fixed function of the class X_μ. Then

$$
\begin{aligned}
\int_0^{2\pi} S(u)h(u)\,du &= \int_0^{2\pi} h(u)\left(\int\int_{\Omega_u} |f'(z)|^2\,dx\,dy\right)du \\
(15) \qquad &= \int_0^{2\pi} h(u)\left(\int\int_{\Omega_u'} |f'(z)|^2\,dx\,dy\right)du \\
&\quad + \int_0^{2\pi} h(u)\left(\int\int_{\Omega_u''} f|'(z)|^2\,dx\,dy\right)du = P + Q
\end{aligned}
$$

say, where P and Q denote the parts of Ω_u situated inside and outside the circle $|z| = 1 - \epsilon$, respectively (cf. condition (ii) above).

We first shall evaluate the integral Q. Let us denote for this purpose by $\pm\xi_\rho$, where $1 - \epsilon \leqq \rho < 1$, the arguments of the points where the circle $|z| = \rho$ meets the directions

$$\arg(1 - z) = \pm(\tfrac{1}{2}\pi - \delta).$$

It is easy to see that

$$(16) \qquad \xi_\rho \leqq K(1 - \rho),$$

where $K = K_\delta$ depends only on δ. Let, moreover, $\chi(u, \delta)$, $\delta < \pi$, denote the characteristic function of the interval $(-\delta, \delta)$, that is, the function equal to 1 for $-\delta \leqq u \leqq \delta$ and equal to 0 otherwise (mod 2π). We may write

$$
\begin{aligned}
Q &= \int_0^{2\pi} h(u)\left(\int\int_{\Omega_u''} |f'(\rho e^{i\theta})|^2 \rho\,d\rho\,d\theta\right) \\
&\leqq \int_0^{2\pi} h(u)\left(\int_{1-\epsilon}^1 \rho\,d\rho \int_{u-\xi_\rho}^{u+\xi_\rho} |f'(\rho e^{i\theta})|^2\,d\theta\right)du \\
&= \int_0^{2\pi} h(u)\left(\int_{1-\epsilon}^1 \rho\,d\rho \int_0^{2\pi} |f'(\rho e^{i\theta})|^2 \chi(\theta - u, \xi_\rho)\,d\theta\right)du \\
&= \int_{1-\epsilon}^1 \rho\,d\rho \int_0^{2\pi} \xi_\rho |f'(\rho e^{i\theta})|^2\,d\theta \left(\frac{1}{\xi_\rho}\int_0^{2\pi} h(u)\chi(u - \theta, \xi_\rho)\,du\right).
\end{aligned}
$$

Hence, using Lemma 3, we obtain

$$
\begin{aligned}
Q &\leqq 2\int_{1-\epsilon}^1 \rho\,d\rho \int_0^{2\pi} \xi_\rho |f'(\rho e^{i\theta})|^2 h^*(\theta)\,d\theta \\
&\leqq 2K \int_{1-\epsilon}^1 \rho\,d\rho \int_0^{2\pi} (1 - \rho)|f'(\rho e^{i\theta})|^2 h^*(\theta)\,d\theta \\
&\leqq 2K \int_0^{2\pi} h^*(\theta)\left(\int_0^1 (1 - \rho)|f'(\rho e^{i\theta})|^2\,d\rho\right)d\theta \\
&= 2K \int_0^{2\pi} h^*(\theta) g^2(\theta)\,d\theta.
\end{aligned}
$$

To the last integral we apply in succession Hölder's inequality, Lemma 3, inequality (14) and Lemma 1. This gives

$$(17)\qquad \begin{aligned} Q &\leqq 2K\left(\int_0^{2\pi} \{h^*(\theta)\}^\mu \, d\theta\right)^{1/\mu}\left(\int_0^{2\pi} g^\lambda(\theta)\, d\theta\right)^{2/\lambda} \\ &\leqq 4KB_\lambda^2\left(\frac{\mu}{\mu-1}\right)\left(\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{2/\lambda}. \end{aligned}$$

As regards the integral P, its definition and inequality (14) give

$$(18)\qquad P \leqq \pi \max_{|z| \leqq 1-\epsilon} | f'(z) |^2 \cdot \int_0^{2\pi} h(u)\, du \leqq \pi(2\pi)^{1-1/\mu} \cdot \max_{|z| \leqq 1-\epsilon} | f'(z) |^2.$$

On the other hand, for $| z | \leqq 1 - \epsilon$ we have (cf. (5))

$$\begin{aligned} | f'(z) |^2 &= \left| \frac{1}{2\pi i} \int_{|\zeta| = 1 - \frac{1}{2}\epsilon} \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta \right| \leqq \frac{2}{\pi\epsilon^2} \int_{|\zeta| = 1 - \frac{1}{2}\epsilon} | f(\zeta) | \, d\theta \\ &\leqq \frac{4}{\epsilon^2} I_\lambda^{1/\lambda}(1 - \tfrac{1}{2}\epsilon) \leqq \frac{4}{\epsilon^2}\left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{1/\lambda}. \end{aligned}$$

Substituting this into (18) and taking account of (12), we obtain

$$(19)\qquad P \leqq \frac{4\pi}{\epsilon^2}\left(\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{2/\lambda}.$$

From (13), (15), (17), and (19) we obtain (8). The latter is therefore established in the case $\lambda > 2$, provided the condition (9) is satisfied.

4. That the condition (9) is superfluous may be seen from the following remarks. Since the function $f(z) - a_0$ vanishes at the origin, and since $s(u, f - a_0) = s(u, f)$, inequality (8) is certainly true for $\lambda > 2$, provided we replace $f(z)$ on the right side by $f(z) - a_0$. Minkowski's inequality gives

$$\begin{aligned} \left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) - a_0 |^\lambda \, d\theta\right)^{1/\lambda} &\leqq \left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{1/\lambda} + | a_0 | \\ &\leqq \left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{1/\lambda} + \frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) | \, d\theta \leqq 2\left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{1/\lambda}. \end{aligned}$$

Hence, in the case $f(0) \neq 0$, the constant $A_{\lambda,\Gamma}$ in (8) is merely multiplied by 2.

We shall now extend our theorem to the general case $\lambda > 0$. Let K^λ denote the upper bound of the integral (2) for $0 \leqq r < 1$. That is,

$$\left(\frac{1}{2\pi}\int_0^{2\pi} | f(e^{i\theta}) |^\lambda \, d\theta\right)^{1/\lambda} = K.$$

It is known[5] that then

$$f(z) = f_1(z) + f_2(z),$$

[5] See F. Riesz, loc. cit., or Hardy and Littlewood, *Some new properties of Fourier constants*, Math. Annalen, vol. 97(1926), pp. 159–209.

where the functions f_1 and f_2 are without zeros in $|z| < 1$, and

$$I_\lambda(r, f_1) \leqq (2K)^\lambda, \qquad I_\lambda(r, f_2) \leqq (2K)^\lambda, \qquad (0 \leqq r \leqq 1).$$

Since, by Minkowski's theorem,

$$s(u, f) \leqq s(u, f_1) + s(u, f_2),$$

it is sufficient to prove Theorem 1 for f_1 and f_2. In other words, we may assume that the function f has no zeros for $|z| < 1$.

It should also be observed that the value of the constant $A_{\lambda,\Gamma}$ which we obtained for $\lambda > 2$ tends to ∞ as λ tends to 2 (cf. (7)). For $\lambda \geqq \lambda_0 > 2$ that constant is bounded. The argument which follows will incidentally show that the least value which we may take for $A_{\lambda,\Gamma}$ is bounded for $\lambda \geqq \lambda_0 > 0$. We may restrict our considerations to the case $0 < \lambda \leqq 3$.

We write[6]

$$F(z) = \{f(z)\}^{\frac{1}{4}\lambda}. \tag{20}$$

The function $F(z)$ belongs to H^4. Hence

$$\left(\int_0^{2\pi} s^4(u, F)\, du\right)^{\frac{1}{4}} \leqq A_{4,\Gamma}\left(\int_0^{2\pi} |F(e^{i\theta})|^4\, d\theta\right)^{\frac{1}{4}}. \tag{21}$$

On the other hand, $f(z) = \{F(z)\}^{4/\lambda}$, and so

$$f'(z) = \frac{4}{\lambda}\{F(z)\}^{4/\lambda - 1}F'(z),$$

$$s(u, f) \leqq \frac{4}{\lambda}\{M(u, F)\}^{4/\lambda-1}s(u, F)$$

(cf. Lemma 2). An application of Hölder's inequality gives

$$\begin{aligned}\int_0^{2\pi} s^\lambda(u, f)\, du &\leqq \left(\frac{4}{\lambda}\right)^\lambda \int_0^{2\pi} s^\lambda(u, F)M^{4-\lambda}(u, F)\, du \\ &\leqq \left(\frac{4}{\lambda}\right)^\lambda \left(\int_0^{2\pi} s^4(u, F)\, du\right)^{\frac{1}{4}\lambda}\left(\int_0^{2\pi} M^4(u, F)\, du\right)^{\frac{1}{4}(4-\lambda)}.\end{aligned}$$

Hence, in view of (21), of Lemma 2, and of (20), we obtain

$$\begin{aligned}\left(\int_0^{2\pi} s^\lambda(u, f)\, du\right)^{1/\lambda} &\leqq \frac{4}{\lambda} A_{4,\Gamma} C_{4,\Gamma}^{4/\lambda-1}\left(\int_0^{2\pi} |F(e^{i\theta})|^4\, d\theta\right)^{1/\lambda} \\ &= \frac{4}{\lambda} A_{\lambda,\Gamma} C_{4,\Gamma}^{4/\lambda-1}\left(\int_0^{2\pi} |f(e^{i\theta})|^\lambda\, d\theta\right)^{1/\lambda}.\end{aligned}$$

This completes the proof of Theorem 1. When $0 < \lambda \leqq 3$ we may take

$$A_{\lambda,\Gamma} = \frac{4}{\lambda} A_{4,\Gamma} C_{4,\Gamma}^{4/\lambda-1}.$$

[6] The argument is modeled on a similar argument of Littlewood and Paley, loc. cit., p. 69.

5. It may be of some interest to investigate the relations between the functions $s(u)$ and $g(u)$. If, for example, we had an inequality

$$s(u) \leqq \alpha g(u), \tag{22}$$

where α would denote a constant (depending possibly on Γ), Theorem 1 would then be a simple corollary of Lemma 1. The inequality (22), however, is not true. A simple example is given by the function

$$f(z) = e^{-(1-z)^{-4}},$$

for which $g(0)$ is finite, and $s(0)$ is infinite, if only the angle between the curve Γ and the circle $|z| = 1$ at the point $z = 1$ is less than $\frac{1}{4}\pi$. On the contrary, the theorem which follows shows that the inequality opposite to (22) holds for every u.

THEOREM 2. *Let Ω denote the interior of a curve Γ satisfying, besides conditions* (i), (ii), (iv) *stated above, the following condition*:

(v) *Γ meets the real axis at the point $z = 1$ at an angle greater than* 0.

Then,

$$g(u) \leqq \beta s(u), \tag{23}$$

where $\beta = \beta(\Gamma)$ depends only on Γ.

This theorem is not very deep. It is sufficient to prove it for $u = 0$. Let us consider the values of the function $|f'(\rho)|$ on the interval $1 - 2^{-n} \leqq \rho \leqq 1 - 2^{-n-1}$ ($n = 0, 1, 2, \cdots$), and let r_n denote a point of that interval where $|f'(\rho)|$ attains its maximum. Plainly

$$g^2(0) \leqq \sum_{n=0}^{\infty} |f'(r_n)|^2 \int_{1-2^{-n}}^{1-2^{-n-1}} (1-\rho)\,d\rho \leqq \sum_{n=0}^{\infty} \left|\frac{f'(r_n)}{2^n}\right|^2. \tag{24}$$

We now observe that

$$|f'(r)| = \left|\frac{1}{2\pi i}\int_{|\zeta - r| = \rho} \frac{f'(\zeta)}{\zeta - r}\,d\zeta\right|,$$

and so

$$|f'(r)|^2 \leqq \frac{1}{2\pi}\int_0^{2\pi} |f'(r + \rho e^{i\theta})|^2\,d\theta. \tag{25}$$

Let $C(r, R)$ denote the circle $|z - r| \leqq R$, where $0 \leqq r < r + \rho < 1$. We suppose R so small that $C(r, R)$ lies in $|z| < 1$. We multiply the inequality (25) by ρ and integrate over the interval $0 \leqq \rho \leqq R$. We then obtain

$$\frac{1}{2}R^2 |f'(r)|^2 \leqq \frac{1}{2\pi}\int\int_{C(r,R)} |f'(z)|^2\,dx\,dy. \tag{26}$$

Here we substitute $r = r_n$, $R = \eta 2^{-n}$, where η is a constant. If we use condition (v), it is not difficult to see that, if η is small enough, all the circles $C(r_{2k}, \eta 2^{-2k})$ are contained in Ω, and no two of them have points in common. Hence

$$\pi\eta^2 \sum_{k=0}^{\infty} \left|\frac{f'(r_{2k})}{2^{2k}}\right|^2 \leqq \int\int_{\Omega} |f'(z)|^2\,dx\,du. \tag{27}$$

This inequality holds if we replace the even indices $2k$ by the odd indices $2k + 1$. Adding the new inequality to (27) and taking account of (24), we obtain (23).

6. In the case $\lambda > 1$, the inequality opposite to (8) is also true. More precisely, we have

THEOREM 3. *If $f(z)$ is of H^λ, $\lambda > 1$, if $f(0) = 0$, and if Γ satisfies conditions* (i)-(v), *then*

$$\left\{\int_0^{2\pi} |f(e^{i\theta})|^\lambda d\theta\right\}^{1/\lambda} \leqq A_{\lambda,\Gamma}\left\{\int_0^{2\pi} s^\lambda(u)\, du\right\}^{1/\lambda},$$

where $A_{\lambda,\Gamma}$ depends only on λ and Γ.

This result is an immediate consequence of Theorem 2 and of the following known proposition.

LEMMA 4. *The inequality opposite to* (11), *that is, the inequality*

$$\left\{\int_0^{2\pi} |f(e^{i\theta})|^\lambda d\theta\right\}^{1/\lambda} \leqq B_\lambda\left\{\int_0^{2\pi} g^\lambda(u)\, du\right\}^{1/\lambda}$$

is true for $\lambda > 1$, provided $f(0) = 0$.[7]

Part II

7. THEOREM 4. *Let $f(z)$ be a function holomorphic in the circle $|z| < 1$, and let E be a set of positive measure situated on the circumference $|z| = 1$ and having the following property:*

The limit

$$f(e^{iu}) = \lim_{z\to e^{iu}} f(z)$$

exists and is finite when z tends to any point e^{iu} of E along any non-tangential path.

Then, for almost every point of E, the integral (7) *is finite.*

Let Δ denote the curvilinear triangle

$$|\arg(1 - z)| \leqq \tfrac{1}{4}\pi, \qquad |z| \geqq 2^{-\frac{1}{2}}, \qquad |\arg z| \leqq \tfrac{1}{4}\pi.$$

By Δ_u we shall mean the triangle Δ rotated about $z = 0$ through an angle u. To every point e^{iu} belonging to E, there corresponds a number N_u such that

$$|f(z)| \leqq N_u \qquad \textit{for } z \textit{ belonging to } \Delta_u.$$

There exists a perfect subset P of E and a number N such that

$$N_u \leqq N \qquad \textit{for } e^{iu} \textit{ belonging to } P.$$

The measure of P may differ as little as we please from that of E.

Let R denote the sum of the circle $|z| \leqq \frac{1}{2}$ and of all the triangles Δ_u, for e^{iu} belonging to P. It is not difficult to see that R is a closed set (of a star-like

[7] Littlewood and Paley, loc. cit.

shape), whose boundary is a simple Jordan curve J. The only points which J has in common with the circumference $|z| = 1$ are those of P. What is important is that the curve J is rectifiable.[8]

Let $z = x + iy$, $\zeta = \xi + i\eta$, and let

$$z = \varphi(\zeta) \tag{28}$$

be the function mapping conformally the unit circle $|\zeta| < 1$ into the interior of R. Let

$$\zeta = \psi(z) \tag{29}$$

be the inverse transformation, defined for z belonging to the interior of R. It is well known that the one-to-one transformation defined by (28) and (29) may be extended in a continuous manner to the boundaries $|\zeta| = 1$ and J. The function

$$\Phi(\zeta) = f(\varphi(\zeta))$$

is plainly bounded for $|\zeta| < 1$.

Let W^* denote the point set analogous to Ω, that is, let the boundary C^* of W^* satisfy conditions (i), (iii), (iv) of §1. We shall suppose, instead of condition (ii), that C^* possesses at the point $\zeta = 1$ two one-sided tangents, and that each of these half-tangents makes with the circumference $|\zeta| = 1$ an angle $\frac{1}{2}\delta > 0$. By W_θ^* we shall denote the set W^* rotated about $\zeta = 0$ through an angle θ, and by W_θ the image of W_θ^* defined by the function (28) and situated in R.

In view of Lusin's Theorem A, the integral

$$\iint_{W_\theta^*} |\Phi'(\zeta)|^2 \, d\xi \, d\eta \tag{30}$$

is finite for almost every θ. We may write

$$\iint_{W_\theta^*} |\Phi'(\zeta)|^2 \, d\xi \, d\eta = \iint_{W_\theta} \left| \frac{d}{dz} \Phi(\zeta) \cdot \frac{dz}{d\zeta} \right|^2 \cdot \left| \frac{d\zeta}{dz} \right|^2 dx \, dy = \iint_{W_\theta} |f'(z)|^2 \, dx \, dy. \tag{31}$$

In the sequel, we shall use the following well-known facts concerning the conformal mapping of the circle $|\zeta| < 1$ into the interior of a simple closed curve J of finite length.

(a) *The sets of measure zero situated on the circumference* $|\zeta| = 1$ *are transformed into sets of zero length on* J, *and conversely, the sets of zero length situated on* J *are transformed into sets of measure zero on* $|\zeta| = 1$.[9]

[8] The argument by means of which we deduce Theorem 4 from Lusin's Theorem A is known and has already been used in similar problems. It seems, however, that Theorem 4 has never been stated explicitly.

[9] See F. and M. Riesz, *Über Randwerte einer analytischen Funktion*, Quatrième Congrès des mathématiciens scandinaves, 1916, pp. 27–44. The result was found independently by N. Lusin. Cf. Privaloff, *The Integral of Cauchy* (in Russian), Saratoff, 1919, pp. 1–94.

(b) *At almost every point $e^{i\theta}$, the transformation preserves angles, that is, if C^* is any curve approaching the point $\zeta = e^{i\theta}$, and making an angle α with the circumference $|\zeta| = 1$, the angle which the image C of C^* makes with the tangent to J is also equal to α.*[10]

Let us now assume that Theorem 4 is false, that is, that the integral (7) is infinite for e^{iu} belonging to a subset of E of positive measure. Without loss of generality, we may assume that this subset coincides with P, so that $S(u) = \infty$ for e^{iu} belonging to P.

The set P is situated on the curve J and is of positive length. Let P^* be the image of P on the circumference $|\zeta| = 1$. In view of (a), P^* is of positive measure.

Let P_0^* be a subset of P^* of measure zero and such that, for $e^{i\theta}$ belonging to $P^* - P_0^*$,

(A) the integral (30) is finite;

(B) condition (b) is satisfied;

(C) at the point on J corresponding to $e^{i\theta}$, the tangent to J exists and coincides with the tangent to the circumference $|z| = 1$.

The set P_0^* is of measure zero, so that the measure of $P^* - P_0^*$ is positive. Correspondingly, if P_0 is the image of P_0^* on J, the length of $P - P_0$ is positive.

Let θ be any number such that $e^{i\theta}$ belongs to $P^* - P_0^*$, and let

$$e^{iu} = \varphi(e^{i\theta}).$$

In view of conditions (B) and (C), the frontier C_u of W_θ has also half-tangents at the point e^{iu} and the angles between those half-tangents and the circumference $|z| = 1$ are also equal to $\frac{1}{2}\delta$. In particular, in the neighborhood of the point e^{iu} the set Ω_u is contained in W_θ (cf. condition (ii) of §1). Since the integral (31) is finite, $S(u)$ must also be finite, contrary to assumption. This contradiction proves the theorem.[11]

The theorem holds in the case when $f(z)$ is not holomorphic, but meromorphic in the circle $|z| < 1$. The proof undergoes but little change.

8. As an immediate corollary of Theorems 2 and 3, we obtain

THEOREM 5. *Under the hypothesis of Theorem 4, the function $g(u)$ defined by* (10) *is finite for almost every e^{iu} belonging to E.*

Part III

9. By means of Theorems 1 and 2, we may obtain generalizations of certain inequalities of Littlewood and Paley.

Let r, p, q denote real numbers satisfying the conditions

$$r > 1, \qquad 1 < p \leqq 2 \leqq q. \tag{32}$$

[10] Privaloff, loc. cit., p. 36.

[11] Starting from Theorem 4, as it is actually established, it is possible to show that it holds for *some* curves Γ tangent to the circle $|z| = 1$ at the point $z = 1$. Cf. also Lusin, loc. cit., p. 141.

Let $F(\theta)$ be a function of period 2π and of the class L^r. We shall suppose throughout that the mean value of $F(\theta)$ over the interval $(0, 2\pi)$ is equal to 0. By $\phi(z) = \phi(z, F)$ we shall mean the analytic function whose real part is the Poisson integral of F and which vanishes at the point $z = 0$. We write

$$J_r(\theta) = J_r(\theta, F) = \left(\int_0^1 (1-\rho)^{r-1} \mid \phi'(\rho e^{i\theta}) \mid^r d\theta\right)^{1/r}$$

so that $J_2(\theta) = g(\theta)$. Littlewood and Paley have shown that

$$\int_0^{2\pi} J_q^q(\theta)\, d\theta \leqq A_q^q \int_0^{2\pi} \mid F(\theta) \mid^q d\theta, \tag{33}$$

$$\int_0^{2\pi} \mid F(\theta) \mid^p d\theta \leqq A_p^p \int_0^{2\pi} J_p^p(\theta)\, d\theta, \tag{34}$$

where A_r depends only on r.[12] We shall prove the following proposition which contains (33) and (34) as special cases.

THEOREM 6. *If the numbers p, q, r satisfy (32), and if $F(\theta)$ belongs to L^r, then*

$$\left(\int_0^{2\pi} J_q^r(\theta)\, d\theta\right)^{1/r} \leqq K_r \left(\int_0^{2\pi} \mid F(\theta) \mid^r d\theta\right)^{1/r}, \tag{35}$$

$$\left(\int_0^{2\pi} \mid F(\theta) \mid^r d\theta\right)^{1/r} \leqq L_r \left(\int_0^{2\pi} J_p^r(\theta)\, d\theta\right)^{1/r}, \tag{36}$$

where K_r and L_r depend only on r.

We begin by proving (35). Let

$$\mu_n(\theta) = \max_{1-2^{-n} \leqq \rho \leqq 1-2^{-n-1}} \mid \phi'(\rho e^{i\theta}) \mid,$$

$$g^*(\theta) = \left(\sum_{n=0}^{\infty} \frac{\mu_n^2(\theta)}{2^{2n}}\right)^{\frac{1}{2}}.$$

Let $s(u)$ have the meaning of Theorem 2. The formula (24) may be written as $g(\theta) \leqq g^*(\theta)$, and from the proof of Theorem 2 it follows that

$$g^*(\theta) \leqq \beta s(\theta). \tag{37}$$

If we fix the curve Γ, the constant β of this inequality is an absolute constant. From (37) and (8) it follows that

$$\int_0^{2\pi} \left(\sum_{n=0}^{\infty} \frac{\mu_n^2(\theta)}{2^{2n}}\right)^{\frac{1}{2}r} d\theta \leqq A_r^r \int_0^{2\pi} \mid f(e^{i\theta}) \mid^r d\theta \leqq C_r^r \int_0^{2\pi} \mid F(\theta) \mid^r d\theta, \tag{38}$$

where A_r and C_r depend only on r. (The last inequality is a consequence of the well-known M. Riesz inequality concerning conjugate functions.) On the other hand,

$$\left(\sum_{n=0}^{\infty} \frac{\mu_n^2(\theta)}{2^{2n}}\right)^{\frac{1}{2}r} \geqq \left(\sum_{n=0}^{\infty} \frac{\mu_n^q(\theta)}{2^{qn}}\right)^{r/q} \geqq J_q^r(\theta). \tag{39}$$

The inequality (35) follows from (38) and (39).

[12] Littlewood and Paley, loc. cit., p. 54.

10. The inequality (36) requires a different treatment. It is, in a sense, a consequence of the inequality (35), but the proof of this fact is by no means trivial. The argument which follows is, with slight changes, a repetition of the argument by which Littlewood and Paley deduced (34) from (33).

We observe first of all that it is sufficient to prove (36) in the case when $\phi(z)$ is holomorphic in the closed circle $|z| \leqq 1$. For otherwise, instead of $F(\theta)$ and $\phi(z)$, we may consider $F_R(\theta) = U(R, \theta)$ and $\phi(Rz)$, where $U(R, \theta)$ denotes the Poisson integral of $F(\theta)$, and $R < 1$. It is easy to verify that $J_p(\theta, F_R) \leqq J_p(\theta, F)$. Hence, if (36) is true in the special case just mentioned, we may write

$$\left(\int_0^{2\pi} |U(R,\theta)|^r d\theta\right)^{1/r} \leqq L_r \left(\int_0^{2\pi} J_p^r(\theta, F_R)\, d\theta\right)^{1/r} \leqq L_r \left(\int_0^{2\pi} J_p^r(\theta, F)\, d\theta\right)^{1/r}.$$

Hence the inequality (36) follows on making $R \to 1$.

We shall require the following lemma.

LEMMA 5. *Let $\phi(z)$ be holomorphic in the circle $|z| \leqq 1$, and let $H(\theta)$ be defined by the equation*

$$H(\theta) = |F(\theta)|^{r-1}\, \overline{\operatorname{sign} F(\theta)} + c\left(\frac{1}{2\pi}\int_0^{2\pi} |F(\theta)|^r d\theta\right)^{(r-1)/r}, \tag{40}$$

the constant c being chosen to make $\int_0^{2\pi} H(\theta)\, d\theta = 0$. If $\psi(z)$ is the function ϕ corresponding to $H(\theta)$, then

$$\begin{aligned}&\int_0^{2\pi} |F(\theta)|^r d\theta \\ &\leqq M_r^r \int_0^{2\pi}\int_0^1 (1-\rho)\left\{|\phi'(\rho e^{i\theta})\psi'(\rho e^{i\theta})| + \frac{1}{\rho}|\phi'(\rho e^{i\theta})\psi(\rho e^{i\theta})|\right\} d\theta,\end{aligned} \tag{41}$$

where M_r depends only on r.[13]

From the definition of c, we deduce that $|c| \leqq 1$, and so

$$\int_0^{2\pi} |H(\theta)|^{r'} d\theta \leqq 2^{r'} \int_0^{2\pi} |F(\theta)|^r d\theta, \tag{42}$$

where $r' = r/(r-1)$ is the exponent conjugate to r. Let $J_r^*(\theta)$ denote the function $J_r(\theta)$ corresponding to H, and let

$$\Psi(\theta) = \max_{0 \leqq \rho < 1} \rho^{-1} |\psi(\rho e^{i\theta})|.$$

We may write

$$\begin{aligned}&\int_0^{2\pi}\int_0^1 (1-\rho)\,|\phi'(\rho e^{i\theta})\psi'(\rho e^{i\theta})|\, d\rho\, d\theta \\ &\leqq \int_0^{2\pi} d\theta \left(\int_0^1 (1-\rho)^{p-1} |\phi'(\rho e^{i\theta})|^p d\rho\right)^{1/p} \left(\int_0^1 (1-\rho)^{p'-1} |\psi'(\rho e^{i\theta})|^{p'} d\rho\right)^{1/p'} \\ &\leqq \int_0^{2\pi} J_p(\theta) J_{p'}^*(\theta)\, d\theta \leqq \left(\int_0^{2\pi} J_p^r(\theta)\, d\theta\right)^{1/r} \left(\int_0^{2\pi} J_{p'}^{*r'}(\theta)\, d\theta\right)^{1/r'}.\end{aligned} \tag{43}$$

[13] Littlewood and Paley, loc. cit., p. 71.

Since $p' > 2$, inequalities (33) and (42) give

$$(44)\quad \left(\int_0^{2\pi} J_{p'}^{*r'}(\theta)\,d\theta\right)^{1/r'} \leqq K_{r'}\left(\int_0^{2\pi} |H(\theta)|^{r'}\,d\theta\right)^{1/r'} \leqq 2K_{r'}\left(\int_0^{2\pi} |F(\theta)|^{r}\,d\theta\right)^{1/r'}.$$

Similarly,

$$\int_0^{2\pi}\int_0^1 (1-\rho)\rho^{-1}|\phi'(\rho e^{i\theta})\psi(\rho e^{i\theta})|\,d\rho\,d\theta \leqq \int_0^{2\pi}\Psi(\theta)\int_0^1 (1-\rho)|\phi'(\rho e^{i\theta})|\,d\rho\,d\theta$$

$$\leqq \int_0^{2\pi}\Psi(\theta)\left(\int_0^1 (1-\rho)^{p-1}|\phi'(\rho e^{i\theta})|^p\,d\rho\right)^{1/p}\left(\int_0^1 (1-\rho)^{p'-1}\,d\rho\right)^{1/p'}$$

$$\leqq \int_0^{2\pi}\Psi(\theta)J_p(\theta)\,d\theta \leqq \left(\int_0^{2\pi} J_p^r(\theta)\,d\theta\right)^{1/r}\left(\int_0^{2\pi}\Psi^{r'}(\theta)\,d\theta\right)^{1/r'}$$

$$\leqq C_{r'}\left(\int_0^{1} J_p^r(\theta)\,d\theta\right)^{1/r}\left(\int_0^{2\pi}|H(\theta)|^{r'}\,d\theta\right)^{1/r'}$$

on account of Lemma 2.[14] Hence, in view of (42),

$$(45)\quad \begin{aligned}&\int_0^{2\pi}\int_0^1 (1-\rho)\rho^{-1}|\phi'(\rho e^{i\theta})\psi(\rho e^{i\theta})|\,d\rho\,d\theta\\ &\qquad\leqq 2C_{r'}\left(\int_0^{2\pi} J_p^r(\theta)\,d\theta\right)^{1/r}\left(\int_0^{2\pi}|F(\theta)|^r\,d\theta\right)^{1/r'}.\end{aligned}$$

From (41), (43), (44), and (45) we obtain

$$\int_0^{2\pi}|F(\theta)|^r\,d\theta \leqq 2M_r^r(C_{r'}+K_{r'})\left(\int_0^{2\pi} J_p^r(\theta)\,d\theta\right)^{1/r}\left(\int_0^{2\pi}|F(\theta)|^r\,d\theta\right)^{1/r'},$$

and so

$$\left(\int_0^{2\pi}|F(\theta)|^r\,d\theta\right)^{1/r} \leqq 2M_r(C_{r'}+K_{r'})^{1/r}\left(\int_0^{2\pi} J_p^r(\theta)\,d\theta\right)^{1/r} = L_r\left(\int_0^{2\pi} J_p^r(\theta)\,d\theta\right)^{1/r}.$$

This completes the proof of (34).

Remark. It is not difficult to see that if we write inequality (35) in the form

$$(46)\quad \left(\int_0^{2\pi} J_q^r(\theta)\,d\theta\right)^{1/r} \leqq K_r\left(\int_0^{2\pi}|\phi(e^{i\theta})|^r\,d\theta\right)^{1/r},$$

then the inequality remains valid for any $r > 0$. The condition $\phi(0) = 0$ is superfluous both for (35) and for (46). For the validity of (36), this condition is plainly indispensable.

UNIVERSITY OF WILNO.

[14] We observe that $|\psi(z)/z|$ does not attain its maximum in the circle $|z| < \frac{1}{2}$, and for $|z| \geqq \frac{1}{2}$ we have $|\psi(z)/z| \leqq 2|\psi(z)|$.

SUBRINGS OF INFINITE DIRECT SUMS

By Neal H. McCoy

1. **Introduction.** In an attempt to find, for arbitrary rings, analogues of the ordinary direct sum decompositions for rings in which a chain condition is satisfied, one is led at once to the necessity of generalizing the notion of direct sum. If we have a set of rings S_α ($\alpha \subset \mathfrak{A}$), where $\mathfrak{A}$ is an arbitrary set of indices, we shall understand the *direct sum* of the rings S_α ($\alpha \subset \mathfrak{A}$) to be the ring of all functions defined on $\mathfrak{A}$ such that on α the functional values are in S_α. Otherwise expressed, this direct sum is the set of all formal sums $\sum_{\alpha \subset \mathfrak{A}} a_\alpha$, $a_\alpha \subset S_\alpha$, with addition and multiplication defined as follows:[1]

$$\sum_{\alpha \subset \mathfrak{A}} a_\alpha + \sum_{\alpha \subset \mathfrak{A}} b_\alpha = \sum_{\alpha \subset \mathfrak{A}} (a_\alpha + b_\alpha),$$

$$\Big(\sum_{\alpha \subset \mathfrak{A}} a_\alpha\Big)\Big(\sum_{\alpha \subset \mathfrak{A}} b_\alpha\Big) = \sum_{\alpha \subset \mathfrak{A}} a_\alpha b_\alpha.$$

An interesting application of this notion of direct sum has been made in the study of Boolean rings as defined by Stone.[2] One of Stone's results is equivalent to the statement that every Boolean ring is isomorphic to a subring of a direct sum of two-element Boolean rings, i.e., Galois fields $GF(2)$. This result has also been generalized as follows. Let R_p denote a commutative ring in which for the fixed rational prime p, we have $a^p = a$, $pa = 0$ for every a in R_p, a Boolean ring thus being an R_2. In a joint paper with Montgomery,[3] it was shown that such a ring R_p is isomorphic to a subring of a direct sum of fields $GF(p)$. We note also that Köthe[4] has stated that a commutative regular ring, as defined by von Neumann,[5] is isomorphic to a subring of a direct sum of fields.

Received January 12, 1938.

[1] Cf. G. Köthe, *Die Theorie der Verbände* $\cdots$, Jahresbericht der Deutschen Mathematiker Vereinigung, vol. 47(1937), p. 139.

[2] M. H. Stone, *The theory of representation for Boolean algebras*, Transactions of the American Mathematical Society, vol. 40(1936), pp. 37–111. A *Boolean ring* is a ring with the property that $a^2 = a$ for every element a of R. It follows that R is a commutative ring and that always $a + a = 0$.

[3] N. H. McCoy and Deane Montgomery, *A representation of generalized Boolean rings*, this Journal, vol. 3(1937), pp. 455–459. An extension to an even more general class of rings has also been obtained. See N. H. McCoy, *Subrings of direct sums*, abstracted in the Bulletin of the American Mathematical Society, vol. 43(1937), p. 467; published in full in the American Journal of Mathematics, vol. 60(1938), pp. 374–382.

[4] Köthe, loc. cit., p. 139. See also Köthe, *Abstrakte Theorie nichtkommutativer Ringe* $\cdots$, Mathematische Annalen, vol. 103(1930), pp. 545–572, particularly p. 552.

[5] J. von Neumann, *On regular rings*, Proceedings of the National Academy of Sciences, vol. 22(1936), pp. 707–713. A ring R is *regular* if it has a unit element and if for every element a of R there exists an element x such that $axa = a$. In particular, the Boolean rings with unit element are obviously regular, as are also the generalized Boolean rings R_p.

In §2 of the present paper we point out that the question as to whether a ring R is isomorphic to a subring of a direct sum of rings of some specified kind is equivalent to determining whether the intersection of a certain class of ideals[6] in R is the null ideal. The method is essentially one previously used by Krull,[7] and was discovered independently in the joint paper with Montgomery referred to above. In §3 we apply a theorem of Krull,[8] also previously overlooked in this connection, to obtain almost immediately the general result that *any commutative ring without nilpotent elements is isomorphic to a subring of a direct sum of fields.* This is seen to be an analogue of a well-known theorem to the effect that if the descending chain condition is satisfied,[9] a commutative ring without nilpotent elements is actually a (finite) direct sum of fields.[10] It may be remarked that the analogy is not complete as in general there does not exist an isomorphism with a direct sum of fields, but only with a proper subring of such a direct sum. This is true, even for the special case of Boolean rings, as pointed out by Stone.[11]

It is now natural to inquire whether other direct sum decompositions, valid under "finiteness" conditions, have similar analogues for general rings. It is obvious, for example, that a ring in which the descending chain condition is satisfied is a finite direct sum of irreducible rings. If the chain condition is not satisfied, the situation is not quite so clear. However, we show in §4 that an arbitrary ring, not necessarily commutative, is isomorphic to a *subring* of a direct sum of irreducible rings. Thus, again, we have an analogue of the desired kind.

We may now turn to the known theorem that a commutative ring R, in which both ascending and descending chain conditions are satisfied, is a direct sum of primary rings,[12] and ask whether this theorem has an analogue, of the type discussed above, for the case of an arbitrary commutative ring. In §5 we shall obtain a theorem on ideal arithmetic and then be able to show that such an analogue only exists if R is suitably restricted.

In §6 we make some miscellaneous applications to commutative regular rings. Two different characterizations of these rings are obtained, of which perhaps the following is the more interesting: *A commutative ring R, with unit element,*

[6] Throughout this paper the word "ideal" always means "two-sided ideal". Of course no distinction is necessary in case multiplication is commutative. If an ideal $\mathfrak{b}$ is a divisor of the ideal $\mathfrak{a}$ (that is, all elements of $\mathfrak{a}$ are also elements of $\mathfrak{b}$), we shall sometimes indicate this relation by the notation $\mathfrak{a} \subset \mathfrak{b}$ instead of using the classical notation $\mathfrak{a} \equiv 0\ (\mathfrak{b})$.

[7] W. Krull, *Idealtheorie*, Berlin, 1935, pp. 23–24.

[8] Op. cit., p. 9, or Mathematische Annalen, vol. 101(1929), p. 735.

[9] That is, any sequence of ideals, $\mathfrak{a}_1, \mathfrak{a}_2, \cdots$, where $\mathfrak{a}_i$ is always a proper divisor of $\mathfrak{a}_{i+1}$ can have only a finite number of terms. Similarly, the ascending chain condition requires that any sequence, with $\mathfrak{a}_{i+1}$ a proper divisor of $\mathfrak{a}_i$, has only a finite number of terms.

[10] See, e.g., van der Waerden, *Moderne Algebra*, vol. II, p. 163. References to the first volume of this work will be to the first edition.

[11] Stone, loc. cit., p. 88.

[12] van der Waerden, op. cit., vol. II, p. 163.

is regular if and only if every ideal in R is the intersection of all its prime ideal divisors.

Finally, in an appendix, we sketch an independent proof that *in a commutative ring without nilpotent elements, the intersection of all prime ideals is the null ideal.* This is the part of Krull's structure theorem which is essential for the application in §3.

2. **Subrings of direct sums and intersection of ideals.** The following theorem is fundamental in the study of direct sum decompositions:[13]

THEOREM 1. *Let R be an arbitrary ring, and $\mathfrak{n}_\alpha$ ($\alpha \subset \mathfrak{A}$) a set of ideals in R whose intersection is the null ideal. Then R is isomorphic to a subring of the direct sum of the rings $R/\mathfrak{n}_\alpha$ ($\alpha \subset \mathfrak{A}$).*

Denote the homomorphism $R \to R/\mathfrak{n}_\alpha$ by h_α, and the image of a under this homomorphism by $h_\alpha(a)$. To the element a of R we now make correspond the function

$$y_a(\alpha) \equiv h_\alpha(a) \qquad (\alpha \subset \mathfrak{A}).$$

Since h_α is a homomorphism, it follows at once that

$$y_a(\alpha) + y_b(\alpha) = y_{a+b}(\alpha),$$

and

$$y_a(\alpha)y_b(\alpha) = y_{ab}(\alpha).$$

Thus the set of functions y_a ($a \subset R$) is a ring, and the correspondence $a \to y_a$ defines a ring homomorphism. Furthermore it is actually an isomorphism, for if $a \neq 0$ there exists, by hypothesis, an ideal $\mathfrak{n}_\beta$ not containing a, and thus $h_\beta(a) \neq 0$ so that y_a can not vanish identically. Thus R is isomorphic to a ring of functions defined on $\mathfrak{A}$ such that on α the functional values are in $R/\mathfrak{n}_\alpha$. The proof of the theorem is completed by noting that the set of all such functions is, by definition, the direct sum of the rings $R/\mathfrak{n}_\alpha$ ($\alpha \subset \mathfrak{A}$).

3. **Commutative rings without nilpotent elements.** In this section we shall consider an arbitrary commutative ring R without nilpotent elements. If $\mathfrak{a}$ is an arbitrary ideal in R, the *radical* $\mathfrak{r}$ of $\mathfrak{a}$ is defined to be the ideal consisting of all elements of R of which a power belongs to $\mathfrak{a}$. In particular, if $\mathfrak{a} = (0)$, then $\mathfrak{r} = (0)$ since R has no nilpotent elements. A general structure theorem of Krull[14] then yields at once the following result.

THEOREM 2. *In a commutative ring R without nilpotent elements, the intersection of all prime ideals is the null ideal.*

Another proof of this theorem is to be found in an appendix to the present paper.

We may now apply Theorem 1 as follows. Let $\mathfrak{p}_\alpha$ ($\alpha \subset \mathfrak{A}$) denote the set

[13] Cf. Krull, op. cit., pp. 23–24; also McCoy and Montgomery, loc. cit., p. 455.

[14] Op. cit., p. 9.

of all prime ideals in R. It then follows that R is isomorphic to a subring of a direct sum of rings $R/\mathfrak{p}_\alpha$ ($\alpha \subset \mathfrak{A}$). Now a ring $R/\mathfrak{p}_\alpha$ contains no divisors of zero, and can therefore be imbedded in a field. We thus have the following general result.

THEOREM 3. *A necessary and sufficient condition that a commutative ring R be isomorphic to a subring of a direct sum of fields is that R contain no nilpotent elements.*

We may remark that if R is given, this theorem gives little or no information about the structure of the fields which appear in the direct sum, except in special cases. If, however, R is a generalized Boolean ring R_p as defined above, it is easy to show that the ring $R/\mathfrak{p}_\alpha$ is actually a $GF(p)$ for every α.

4. **Subrings of direct sums of irreducible rings.** We shall now let R be an entirely arbitrary ring, not necessarily commutative. If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in R, we may denote their intersection by $\mathfrak{a} \cap \mathfrak{b}$ and their sum by $\mathfrak{a} \cup \mathfrak{b}$. If $\mathfrak{a}_\beta$ ($\beta \subset \mathfrak{B}$) is an arbitrary, finite or infinite, set of ideals in R, we shall mean by the sum of the $\mathfrak{a}_\beta$ the intersection of all ideals containing all $\mathfrak{a}_\beta$ ($\beta \subset \mathfrak{B}$), and shall denote this sum by $\underset{\beta \subset \mathfrak{B}}{S} \mathfrak{a}_\beta$. It is clear that $\underset{\beta \subset \mathfrak{B}}{S} \mathfrak{a}_\beta$ consists of all finite sums $a_{\beta_1} + a_{\beta_2} + \cdots + a_{\beta_n}$, where $a_{\beta_i} \subset \mathfrak{a}_{\beta_i}$, $\beta_i \subset \mathfrak{B}$.

We may recall that a ring R is said to be *irreducible* if it cannot be expressed as the direct sum of two of its ideals. In this section we shall prove the

THEOREM 4. *Any ring is isomorphic to a subring of a direct sum of irreducible rings.*

In view of the fact that a ring R without unit element can be imbedded in a ring with unit element,[15] there is no loss of generality in assuming henceforth that R has a unit element 1.

An ideal $\mathfrak{a}$ in R is said to be *irreducible* if it cannot be expressed in the form $\mathfrak{a} = \mathfrak{a}_1 \cap \mathfrak{a}_2$, where $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are proper ideal divisors of $\mathfrak{a}$. Also $\mathfrak{a}$ is (direct) *indecomposable* if there do not exist two proper divisors $\mathfrak{a}_1$, $\mathfrak{a}_2$ of $\mathfrak{a}$ such that $\mathfrak{a} = \mathfrak{a}_1 \cap \mathfrak{a}_2$, $\mathfrak{a}_1 \cup \mathfrak{a}_2 = (1)$. It is obvious that an irreducible ideal is also indecomposable. Now we have the known result[16] that if $\mathfrak{a}$ is an ideal in R, the ring $R/\mathfrak{a}$ is irreducible if and only if $\mathfrak{a}$ is indecomposable. Our theorem will

[15] One may, for example, consider the ring S for all pairs (r, n), where r is an element of R and n a rational integer, with addition and multiplication defined as follows:

$$(r_1, n_1) + (r_2, n_2) = (r_1 + r_2, n_1 + n_2),$$

and

$$(r_1, n_1)(r_2, n_2) = (r_1 r_2 + n_2 r_1 + n_1 r_2, r_1 r_2).$$

Then S is seen to be a ring with unit element (0, 1) and the subring of all elements of the form $(r, 0)$ is isomorphic to R. For this construction, see J. L. Dorroh, *Concerning adjunctions to algebras*, Bulletin of the American Mathematical Society, vol. 38(1932), pp. 85–88.

[16] The corresponding theorem for the case of one-sided ideals was proved explicitly by E. Noether and W. Schmeidler in the paper, *Moduln in nichtkommutativen Bereichen*, $\cdots$, Mathematische Zeitschrift, vol. 8(1920), pp. 1–35. The method is readily applied also to two-sided ideals.

then be established by Theorem 1 if we can show that the intersection of all indecomposable ideals in R is merely the ideal (0). This will be established by proving the following

LEMMA. *If a is a fixed element of R other than zero, there exists in R an irreducible, and therefore indecomposable, ideal not containing a.*

It is fairly obvious that the desired irreducible ideal may be constructed by transfinite methods by forming a "largest" ideal not containing a. However, we shall give the construction in some detail as it will also be needed in a different connection in the appendix. It will be noted that the following is an almost trivial adaptation of one of Stone's proofs of the existence of prime ideals in Boolean rings.[17]

Let C be the class of all ideals in R. We assume that the elements of C can be put in a one-to-one correspondence with the ordinals $\gamma < \omega$ for a suitably chosen ordinal ω. The elements of C may thus be denoted by $\mathfrak{a}_\gamma$ $(\gamma < \omega)$ and, in particular, we shall suppose that $\mathfrak{a}_1 = (0)$. We now define a sequence of ideals $\mathfrak{b}_\alpha$ as follows. Let $\mathfrak{b}_1 = \mathfrak{a}_1 = (0)$. If now $\mathfrak{b}_\alpha$ has been defined for all ordinals α such that $\alpha < \beta$ where $\beta < \omega$, we define $\mathfrak{b}_\beta$ to be $\underset{\alpha<\beta}{S}\, \mathfrak{b}_\alpha$ if $a \subset \mathfrak{a}_\beta \cup \underset{\alpha<\beta}{S}\, \mathfrak{b}_\alpha$; otherwise we set $\mathfrak{b}_\beta = \mathfrak{a}_\beta \cup \underset{\alpha<\beta}{S}\, \mathfrak{b}_\alpha$. We now assert that the ideal $\mathfrak{m} = \underset{\alpha<\omega}{S}\, \mathfrak{b}_\alpha$ is an irreducible ideal not containing a, as we proceed to show.

We first show that $\mathfrak{m}$ does not contain a. If $\alpha < \beta$, we note that $\mathfrak{b}_\alpha \subset \mathfrak{b}_\beta$ by definition. If we assume that a is in $\mathfrak{m}$, it follows that $a = b_{\beta_1} + \cdots + b_{\beta_n}$, where $b_{\beta_i} \subset \mathfrak{b}_{\beta_i}$, and $\beta_1 < \beta_2 < \cdots < \beta_n < \omega$. That is, $a \subset \mathfrak{b}_{\beta_n}$, $\beta_n < \omega$. Suppose that β is the first ordinal such that $a \subset \mathfrak{b}_\beta$. If $\beta \neq 1$, it follows from the definition that $\mathfrak{b}_\beta = \underset{\alpha<\beta}{S}\, \mathfrak{b}_\alpha$, and a repetition of the preceding argument shows that $a \subset \mathfrak{b}_\delta$ for some ordinal $\delta < \beta$, a contradiction. Thus, if $a \subset \mathfrak{m}$, $a \subset \mathfrak{b}_1 = \mathfrak{a}_1 = (0)$, which is clearly impossible as we assumed that $a \neq 0$.

Suppose that $\mathfrak{m}$ is reducible, and that $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are proper ideal divisors of $\mathfrak{m}$ such that $\mathfrak{m} = \mathfrak{m}_1 \cap \mathfrak{m}_2$. Then, since $\mathfrak{m}$ does not contain a, at least one of the ideals $\mathfrak{m}_1$, $\mathfrak{m}_2$ does not contain a. Let us assume that $\mathfrak{m}_1$ does not contain a. Suppose by the well-ordering of elements of C that $\mathfrak{m}_1 = \mathfrak{a}_\gamma$. Thus $\mathfrak{m}_1 = \mathfrak{a}_\gamma \supset \mathfrak{m} \supset \underset{\alpha<\gamma}{S}\, \mathfrak{b}_\alpha$. Now, by definition, we must have $\mathfrak{b}_\gamma = \mathfrak{a}_\gamma \cup \underset{\alpha<\gamma}{S}\, \mathfrak{b}_\alpha$ since a is not contained in $\mathfrak{a}_\gamma \cup \underset{\alpha<\gamma}{S}\, \mathfrak{b}_\alpha = \mathfrak{a}_\gamma$. Thus $\mathfrak{m} \supset \mathfrak{b}_\gamma = \mathfrak{a}_\gamma \cup \underset{\alpha<\gamma}{S}\, \mathfrak{b}_\alpha = \mathfrak{a}_\gamma = \mathfrak{m}$, and we have $\mathfrak{m} = \mathfrak{m}_1$, which contradicts the assumption that $\mathfrak{m}_1$ is a proper divisor of $\mathfrak{m}$. The lemma and the theorem are therefore established.

5. **Ideal arithmetic. Subrings of direct sums of primary rings.** Let $\mathfrak{a}$ be an ideal in the arbitrary ring R, and $\bar{\mathfrak{b}}$ an ideal in the quotient ring $R/\mathfrak{a}$. The set of all elements of R whose images are in $\bar{\mathfrak{b}}$ by the homomorphism $R \rightarrow R/\mathfrak{a}$, is an ideal $\mathfrak{b}$ in R which *corresponds* to the ideal $\bar{\mathfrak{b}}$ in R/a. Let P denote some

[17] Stone, loc. cit., p. 101.

property of ideals such that for all choices of R, $\mathfrak{a}$ and $\bar{\mathfrak{b}}$ it is true that if $\bar{\mathfrak{b}}$ has property P, then the corresponding ideal $\mathfrak{b}$ has property P. Since $R/\mathfrak{b} \cong (R/\mathfrak{a})/\bar{\mathfrak{b}}$,[18] it is readily seen that this condition is satisfied in the following cases which will be used in the sequel. (1) If R is commutative, P may be the property of being prime; (2) if R is commutative, P may represent the property of being primary;[19] (3) if R has a unit element, P may stand for indecomposability. An ideal having property P may be said to be of *type* P. We now prove the following theorem:[20]

THEOREM 5. *If for every ideal $\mathfrak{a}$ in a ring R it is true that the intersection of all ideals of type P in the ring $R/\mathfrak{a}$ is the null ideal, then every ideal in R is the intersection of all its divisors of type P.*

We first establish the

LEMMA. *Under the hypotheses of the theorem, let $\mathfrak{c}$ be an ideal in R not containing the ideal $\mathfrak{d}$. Then there exists in R an ideal of type P containing $\mathfrak{c}$ but not $\mathfrak{d}$.*

Let d be an element of $\mathfrak{d}$ not also in $\mathfrak{c}$, and suppose $d \rightarrow \bar{d}$ by the homomorphism $R \rightarrow R/\mathfrak{c}$. Then $\bar{d} \neq 0$, and by hypothesis there exists in $R/\mathfrak{c}$ an ideal $\bar{\mathfrak{m}}$ of type P not containing $\bar{d}$. This implies, by our assumptions on P, that the corresponding ideal $\mathfrak{m}$ in R is also of type P. Clearly $\mathfrak{m}$ contains $\mathfrak{c}$ but not $\mathfrak{d}$, so that $\mathfrak{m}$ satisfies the requirements of the lemma.

Now let $\mathfrak{c}$ be any ideal in R other than the unit ideal and apply the lemma with $\mathfrak{d}$ as the unit ideal. Thus there exists in R at least one ideal of type P containing $\mathfrak{c}$. If $\mathfrak{k}$ denotes the intersection of all ideals of type P which contain $\mathfrak{c}$, then clearly $\mathfrak{k} \supset \mathfrak{c}$. Suppose $\mathfrak{c}$ does not contain $\mathfrak{k}$. Then, if we again apply the lemma, there exists an ideal of type P containing $\mathfrak{c}$ but not $\mathfrak{k}$, contrary to the definition of $\mathfrak{k}$. Thus we must have $\mathfrak{c} \supset \mathfrak{k}$, and therefore $\mathfrak{c} = \mathfrak{k}$ as required.

As a first application of this theorem we let P denote indecomposability and then obtain at once from the lemma of §4 the following

COROLLARY. *In an arbitrary ring with unit element, every ideal is the intersection of all its indecomposable ideal divisors.*

We may now discuss the question as to whether every commutative ring is isomorphic to a subring of a direct sum of primary rings. Suppose this were true, and let R denote an arbitrary commutative ring. Then it is easily seen that if a is an arbitrary element of R other than zero, there exists a homomorphism of R into a subring of a primary ring taking a into an element other than zero. But a subring of a primary ring is also primary, so that there exists

[18] E. Noether, *Hyperkomplexe Grössen und Darstellungstheorie*, Mathematische Zeitschrift, vol. 30(1929), p. 657.

[19] An ideal $\mathfrak{q}$ in a commutative ring R is *primary* if whenever $ab \equiv 0\ (\mathfrak{q})$, $a \not\equiv 0\ (\mathfrak{q})$, then $b^n \equiv 0\ (\mathfrak{q})$ for some positive integer n. Otherwise expressed, this means that in the quotient ring $R/\mathfrak{q}$, every divisor of zero is nilpotent. A ring is primary if the null ideal is primary. Thus a quotient ring $R/\mathfrak{q}$ is primary if and only if $\mathfrak{q}$ is a primary ideal in R.

[20] The proof is a simple generalization of the method used by Stone (loc. cit., p. 105) to establish the "fundamental proposition of ideal arithmetic" for Boolean rings. To avoid trivial special cases we may assume that, in the statement of the theorem, the unit ideal is excluded from consideration.

an ideal $\mathfrak{q}$ in R, not containing a, such that $R/\mathfrak{q}$ is primary. This means, however, that $\mathfrak{q}$ is a primary ideal, and thus the intersection of all primary ideals in R is the null ideal. If, now, this is true for an arbitrary commutative ring, Theorem 5 shows, with P the property of being primary, that every ideal in R is the intersection of all its primary ideal divisors. But it is known[21] that there exist commutative rings in which this is false, and thus not every commutative ring is isomorphic to a subring of a direct sum of primary rings.

If, however, in the commutative ring R, every prime ideal is divisorless,[22] then a result of Krull[23] shows that the intersection of all primary ideals is the null ideal and we have at once from Theorem 1

THEOREM 6. *A commutative ring R in which every prime ideal is divisorless is isomorphic to a subring of a direct sum of primary rings.*

6. **Some applications to commutative regular rings.** A commutative ring R with unit element is *regular* if for every element a of R there exists an element x of R such that $a^2x = a$.[24] It follows that a commutative regular ring has no nilpotent elements. Also a field or a direct sum of fields is seen to be regular. Theorem 3 then yields at once the following result:

THEOREM 7. *Any commutative ring without nilpotent elements can be imbedded in a commutative regular ring.*

If S is any commutative ring with unit element, let us denote by $S(D)$, $S(P)$, $S(I)$, respectively, the classes of divisorless, prime and indecomposable ideals in S. Now $S(D) \subset S(P)$.[25] Also it follows readily that $S(P) \subset S(I)$. For if $\mathfrak{p}$ is a prime ideal in S, $S/\mathfrak{p}$ can not be reducible as it would then contain divisors of zero. Thus $\mathfrak{p}$ is indecomposable and therefore in $S(I)$. We then have the inclusion relations, $S(D) \subset S(P) \subset S(I)$.

THEOREM 8. *For a commutative regular ring R, we have $R(D) = R(P) = R(I)$.*

We need only to show that an indecomposable ideal is divisorless. Let $\mathfrak{a}$ be an arbitrary indecomposable ideal in R. Then $R/\mathfrak{a}$ is an irreducible commutative regular ring and is therefore a field.[26] This implies, however, that $\mathfrak{a}$ is divisorless.

The following theorem gives an interesting characterization of commutative regular rings.

THEOREM 9. *A commutative ring R, with unit element, is regular if and only if each ideal in R is the intersection of all its prime ideal divisors.*

[21] This is the case for certain rings which appear in the general evaluation theory. See W. Krull, *Allgemeine Bewertungstheorie*, Journal für die reine und angewandte Mathematik, vol. 167(1932), p. 167.

[22] That is, has no proper divisor except the unit ideal.

[23] W. Krull, *Idealtheorie in Ringen ohne Endlichkeitsbedingung*, Mathematische Annalen, vol. 101(1929), p. 738.

[24] J. von Neumann, loc. cit.

[25] See van der Waerden, op. cit., vol. I, p. 59.

[26] von Neumann, loc. cit., p. 712.

We first show that a regular ring R has the desired property.[27] If $\mathfrak{a}$ is any ideal in R, then clearly $R/\mathfrak{a}$ is also regular and hence contains no nilpotent elements. Theorem 2 together with Theorem 5 (with P the property of being prime) then yields at once the desired result.

Suppose, now, that R is a commutative ring with unit element, and that every ideal in R is the intersection of all its prime ideal divisors. Let a be any element of R other than zero, and consider the principal ideals (a) and (a^2). Now a prime ideal contains a^2 if and only if it contains a and hence divides (a^2) if and only if it divides (a). Thus, by hypothesis, it follows that $(a) = (a^2)$, and there therefore exists an element x of R such that $a^2x = a$. Hence R is regular.

We may now obtain another characterization of the commutative regular rings as follows.

THEOREM 10. *A commutative regular ring R, with unit element, is regular if and only if $R(P) = R(I)$.*

The necessity of the condition follows from Theorem 8. Let R be a commutative ring, with unit element, such that $R(P) = R(I)$. By the corollary to Theorem 5 it follows that in R each ideal is the intersection of all its prime ideal divisors and thus, by the preceding theorem, R is regular.

Appendix

We shall indicate briefly an independent proof of Theorem 2 along entirely different lines from those used by Krull in the proof of his more general structure theorem.[28] Let R denote an arbitrary commutative ring without nilpotent elements, and a any fixed element of R other than zero. We wish to show the existence in R of a prime ideal $\mathfrak{p}$ not containing a.

If R does not have a unit element, we imbed it in a commutative ring S with unit element in such a way that S contains no nilpotent elements.[29] Now let x be an indeterminate, commutative with elements of S, and consider the ring $S[x]$ of polynomials in x with coefficients in S. In this ring, the principal ideal $\mathfrak{b} = (a^2x - a)$ contains no element of S except the zero. For suppose s is an element of S in $\mathfrak{b}$. Then we have a relation of the type

$$s = (a_0x^n + a_1x^{n-1} + \cdots + a_n)(a^2x - a) \qquad (a_i \subset S).$$

Equating coefficients of x^{n+1} on both sides, we find that $a_0a^2 = 0$. Now $(a_0a)^2 = a(a_0a^2) = 0$, so that we must have $a_0a = 0$ as otherwise a_0a would be a nilpotent element of S. Thus $a_0x^n(a^2x - a) = 0$, and a repetition of the argument shows that $s = 0$. From this result it follows that the ring $T = S[x]/\mathfrak{b}$ contains a

[27] As a matter of fact, this result can also be obtained as a corollary to a theorem of Krull (see footnote 23). It is only necessary to observe that in a commutative regular ring the class of primary ideals coincides with the class of prime ideals. Since also every prime ideal is divisorless, Krull's theorem shows that every ideal in R is the intersection of all its prime ideal divisors.

[28] Krull, *Idealtheorie*, p. 9.

[29] See footnote 15.

subring isomorphic to S which, for convenience, we shall identify with S. Furthermore, in T, there exists an element c (in fact, the polynomial x itself) such that $a^2c = a$. We shall now make use of this fact to prove the existence of a divisorless ideal in T not containing a, and from this our desired result is readily obtained.

Let $\mathfrak{a}_1$ denote the principal ideal $(1 - ac)$ in T. If y is in $\mathfrak{a}_1$, then clearly $ay = 0$. Therefore, since a is neither zero nor nilpotent, a is not an element of $\mathfrak{a}_1$. We now construct the ideal $\mathfrak{m}$ exactly as in the proof of the lemma in §4 with the single exception that we take $\mathfrak{a}_1 = (1 - ac)$ rather than (0). By a repetition of the argument used there, it follows that $\mathfrak{m}$ does not contain a, and that further if $\mathfrak{n}$ is an ideal containing $\mathfrak{m}$ but not a, then $\mathfrak{n} = \mathfrak{m}$. We now show that if $\mathfrak{n}$ contains both $\mathfrak{m}$ and a, then $\mathfrak{n} = (1)$. For, if $\mathfrak{n}$ contains a, it also contains ac, and also $\mathfrak{n} \supset \mathfrak{m} \supset \mathfrak{a}_1 = (1 - ac)$. Therefore $\mathfrak{n} \supset ac + 1 - ac = 1$, and $\mathfrak{n} = (1)$. This shows that $\mathfrak{m}$ is a divisorless ideal in T not containing a.

Since T has a unit element, the divisorless ideal $\mathfrak{m}$ is also prime. But then it follows that $\mathfrak{p} = \mathfrak{m} \cap R$ is a prime ideal in R which does not contain a, and the theorem is established.

INSTITUTE FOR ADVANCED STUDY AND SMITH COLLEGE.

TENSOR PRODUCTS OF ABELIAN GROUPS

By Hassler Whitney

1. **Introduction.** Let G and H be Abelian groups. Their direct sum $G \oplus H$ consists of all pairs (g, h), with $(g, h) + (g', h') = (g + g', h + h')$. If we consider G and H as subgroups of $G \oplus H$, with elements $g = (g, 0)$ and $h = (0, h)$, then we may form $g + h$, and the ordinary laws of *addition* hold. Our object in this paper is to construct a group $G \circ H$ from G and H, in which we can form $g \cdot h$, with the properties of *multiplication*; that is, the distributive laws

$$(1.1) \qquad (g + g') \cdot h = g \cdot h + g' \cdot h, \qquad g \cdot (h + h') = g \cdot h + g \cdot h'$$

hold. Clearly $G \circ H$ must contain elements of the form $\sum g_i \cdot h_i$; it turns out (Theorem 1) that with these elements, *assuming only* (1.1), *we obtain an Abelian group*, which we shall call the *tensor product* of G and H.[1]

The tensor product is known in one important case; namely, in tensor analysis (see §4, (b), and §11), though the definition in the form here given does not seem to have been used. Certain other cases are known (see §4). We refer to the examples there given for further indications of the scope of the theory. A direct product of algebras has been constructed by J. L. Dorroh,[2] by methods closely allied to those of the present paper.

As is to be expected, we see in Part I that when we multiply several groups together, the associative and commutative laws hold; also the distributive laws with respect to direct sums and difference groups. The group of integers plays the rôle of a unit group.[3] The rest of Part I is devoted largely to a study of the relation between groups with operator rings and tensor products; in particular, divisibility properties are considered.

In Part II, a detailed study of tensor products of linear spaces is made; we now assume $rg \cdot h = g \cdot rh$ (r real). With any element α of $G \circ H$ are associated subspaces $G(\alpha)$ of G and $H(\alpha)$ of H; their dimensions equal the minimum number of terms in an expression $\sum g_i \cdot h_i$ for α, and in this expression the g_i and h_i form bases in $G(\alpha)$ and $H(\alpha)$. The elementary operations of tensor algebra are derived, and a direct manner of introducing covariant differentiation is indicated.[4] If the linear spaces are topological, a topology may be introduced into

Received February 23, 1938; presented to the American Mathematical Society, February 26, 1938. See Proceedings of the National Academy of Sciences, vol. 23(1937), p. 290.

[1] This is so even if G and H are not Abelian; see Theorem 11. If G and H are linear or topological, we use a slightly different definition.

[2] J. L. Dorroh, *Concerning the direct product of algebras*, Annals of Mathematics, vol. 36 (1935), pp. 882–885. The author is indebted to the referee for pointing out this paper to him.

[3] In linear spaces, the group of real numbers also is a unit.

[4] Some of these results have been derived independently by H. E. Robbins.

the tensor product. If the spaces are not of finite dimension, there are of course various topologies possible in the product; the one we give is probably at an extreme end, in that a neighborhood of 0 in any topology will contain a neighborhood of the sort here given. The topology has certain defects in that the associative and distributive laws seem not to hold in general with topology preserved. In the case of Hilbert spaces, there is a natural definition of the topology in the product (see Murray and von Neumann, reference in §4, (c)). In the intermediate case of Banach spaces, probably the norm $|\alpha|$ may be defined as the lower bound of numbers $\sum |g_i|\,|h_i|$ for expressions $\sum g_i\cdot h_i$ of α.[5]

In topological groups which contain denumerable dense sets, the product may be given a topology, as is shown in Part III; it agrees with that in Part II when both are defined. Again, in complicated groups, other topologies are possible and perhaps preferable. Finally, for a more complete theory, one must allow infinite sums $\sum g_i\cdot h_i$.

2. **Notations.** Write $G \approx H$ if G and H are isomorphic. The symbol 0 means the zero in any group, or the group with only the zero element. $A \cap B$ is the set of elements in both A and B. ag (a an integer > 0) means $g + \cdots + g$ (a terms); $(-a)g = a(-g)$, $0g = 0$. $g + A$ is the set of all $g + g'$, g' in A; similarly for $A + B$. $g\cdot B$ is the set of all $g\cdot h$, h in B, etc. $aA =$ all ag, g in A. Note that $2A \subset A + A$, etc. Write $a \mid g$ if there is a g' with $ag' = g$; g is then "divisible" by the integer a. $a \mid A$ means $a \mid g$ for all g in A. G is "completely divisible" if for every $a \neq 0$, $a \mid G$, i.e., $aG = G$. The "nullifier" of H in G (of G in H) is the set of all g (all h) such that $g\cdot h = 0$ for all h in H (all g in G).

Let $\sum^* A$ denote the set of all finite sums $a_1 + \cdots + a_k$, a_i in A, any k; this is a subgroup of G (if $A \subset G$). $\sum_i^* A_i$ is the set of all $a_1 + \cdots + a_k$ (a_i in A_i, any k).

Let $G \oplus H$ and $G \ominus G'$ denote direct sums and difference groups. There is a "natural homomorphism" of G into $G \ominus G'$. Some particular groups we shall use are: $I_0 =$ group of integers; $I_\mu = I_0 \ominus \mu I_0 =$ integers mod μ (with elements a_μ for integral a); $Rt =$ rational numbers; $Rl =$ real numbers. Set $G_\mu = G \ominus \mu G$.

I. Discrete groups

3. **Discrete tensor products.** Let G and H be groups (not necessarily Abelian), with the operation $+$. Let $\mathfrak{S}$ be the set of all symbols

$$(g_1, h_1; \cdots; g_n, h_n) \qquad (g_i \text{ in } G,\ h_i \text{ in } H,\ n \text{ any integer } > 0).$$

We add two symbols by the rule

$$(g_1, h_1; \cdots) + (g_{n+1}, h_{n+1}; \cdots) = (g_1, h_1; \cdots; g_{n+1}, h_{n+1}; \cdots).$$

[5] This definition was suggested to me by H. E. Robbins.

Clearly $+$ is associative. We may put any element of $\mathfrak{S}$ in the normal form $(g_1, h_1) + \cdots + (g_n, h_n)$; if we write

$$g_i \times h_i = (g_i, h_i),$$

we obtain

$$(g_1, h_1; \cdots; g_n, h_n) = g_1 \times h_1 + \cdots + g_n \times h_n .$$

Define two equivalence relations as follows:

$$\cdots + (g + g') \times h + \cdots \sim \cdots + g \times h + g' \times h + \cdots, \tag{3.1}$$

$$\cdots + g \times (h + h') + \cdots \sim \cdots + g \times h + g \times h' + \cdots . \tag{3.2}$$

Any succession $s_1 \sim s_2 \sim \cdots \sim s_p$ we shall call an *equivalence sequence* between s_1 and s_p. If two elements s, s' are joined by an equivalence sequence, we say they are *equivalent*, $s \sim s'$. Let also $s \sim s$. The elements of $\mathfrak{S}$ fall into subsets under this relation; these form the elements of the *discrete tensor product* $G \circ H$. In case G and H are discrete, we call this the *tensor product*, in agreement with the definition in Part III. Let $\sum g_i \cdot h_i = g_1 \cdot h_1 + \cdots$ be the element of $G \circ H$ containing the element $\sum g_i \times h_i$ of $\mathfrak{S}$.

To define the group operation, which we temporarily call $\oplus$, in $G \circ H$, take any α and α', and let $\sum g_i \times h_i$ and $\sum g'_i \times h'_i$ be corresponding elements of $\mathfrak{S}$; we set

$$\alpha \oplus \alpha' = \sum g_i \cdot h_i + \sum g'_i \cdot h'_i . \tag{3.3}$$

We must show that this is independent of the choices of $s = \sum g_i \times h_i$ and $s' = \sum g'_i \times h'_i$. If we had chosen t and t', then there are equivalence sequences joining s to t and s' to t'; applying these sequences to $\sum g_i \times h_i + \sum g'_i \times h'_i$ shows that the same element $\alpha \oplus \alpha'$ is determined. Henceforth we use $+$ instead of $\oplus$. Note that $+$ is associative, and (1.1) holds.

We prove in succession the following facts.

(a)
$$\begin{aligned} g\cdot 0 &= (g + g - g)\cdot 0 = g\cdot 0 + g\cdot 0 + (-g)\cdot 0 = g\cdot(0 + 0) + (-g)\cdot 0 \\ &= g\cdot 0 + (-g)\cdot 0 = (g - g)\cdot 0 = 0\cdot 0; \end{aligned}$$

similarly $0\cdot h = 0\cdot 0$.

(b)
$$g\cdot h + 0\cdot 0 = g\cdot h + g\cdot 0 = g\cdot(h + 0) = g\cdot h,$$

and hence $0\cdot 0 = g\cdot 0 = 0\cdot h$ plays the rôle of the identity.

(c)
$$\begin{aligned} g\cdot h &= g\cdot h + 0\cdot(-h) = g\cdot h + g\cdot(-h) + (-g)\cdot(-h) \\ &= g\cdot 0 + (-g)\cdot(-h) = (-g)\cdot(-h). \end{aligned}$$

Next, we may operate with the product as if G and H were Abelian. For

(d)
$$\begin{aligned} g\cdot(h + h') &= g\cdot h + g\cdot h' = (-g)\cdot(-h) + (-g)\cdot(-h') \\ &= (-g)\cdot(-h - h') = g\cdot(h' + h); \end{aligned}$$

similarly $(g + g')\cdot h = (g' + g)\cdot h$. Also

$$g\cdot(h + h' + h'') = g\cdot h + g\cdot(h' + h'') = g\cdot h + g\cdot(h'' + h')$$
$$= g\cdot(h + h'' + h'), \text{ etc.}$$

Finally, the operation in $G \circ H$ is commutative. For [6]

$$\alpha = (g + g')\cdot(h' + h) = g\cdot(h' + h) + g'\cdot(h' + h)$$
$$= g\cdot h' + g\cdot h + g'\cdot h' + g'\cdot h,$$

also

$$\alpha = (g + g')\cdot h' + (g + g')\cdot h = g\cdot h' + g'\cdot h' + g\cdot h + g'\cdot h,$$

and hence

(e) $$g\cdot h + g'\cdot h' = (-g)\cdot h' + \alpha + (-g')\cdot h = g'\cdot h' + g\cdot h.$$

Remark. We would have obtained the same results if we had replaced the elementary equivalence relations by

$$\cdots + (g + g') \times h + \cdots \sim \cdots + g' \times h + g \times h + \cdots, \text{ etc.}$$

THEOREM 1. *$G \circ H$ is an Abelian group; the identity is $0\cdot 0 = g\cdot 0 = 0\cdot h$, and the inverse of $g\cdot h$ is*

(3.4) $$-(g\cdot h) = (-g)\cdot h = g\cdot(-h).$$

The distributive laws (1.1) *hold.*

This follows from the above results. Because of (d), *we henceforth assume G and H are Abelian,* except in Theorem 11.

THEOREM 2. *In any $G \circ H$, for any integer a,*

(3.5) $$a(g\cdot h) = ag\cdot h = g\cdot ah.$$

For instance,

$$(-2)g\cdot h = (-g - g)\cdot h = -[(g + g)\cdot h] = -[g\cdot h + g\cdot h] = (-2)(g\cdot h).$$

Using the distributive laws, we may use summation signs as usual; for instance,

$$\sum_i \Big(\sum_j a_{ij}g_j\Big)\cdot h_i = \sum_i \sum_j (a_{ij}g_j\cdot h_i) = \sum_j \sum_i (g_j\cdot a_{ij}h_i) = \sum_j \Big(g_j\cdot \sum_i a_{ij}h_i\Big).$$

4. **Examples.** A system with both "addition" and "multiplication" may in general be defined by starting with a system or systems, using addition alone,

[6] For a direct proof, we have

$$g\cdot h + g'\cdot h' = g\cdot h + g\cdot h' + (-g + g')\cdot h' = g\cdot(h + h') + (g' - g)\cdot(h + h')$$
$$+ (g' - g)\cdot(-h) = (g + g' - g)\cdot(h + h') + g'\cdot(-h) + (-g)\cdot(-h)$$
$$= g'\cdot(h + h' - h) + g\cdot h = g'\cdot h' + g\cdot h.$$

forming a tensor product, and defining new equality relations. Specifically, any group pair is an example.

(a) The Abelian groups G and H form a *group pair* with respect to the group Z if a multiplication $g \times h = z$ is given, satisfying both distributive laws. *Any such group pair may be defined by choosing a homomorphism of $G \circ H$ into Z.* Clearly

$$\phi(\sum g_i \cdot h_i) = \sum g_i \times h_i$$

has the required properties. Practically all further examples come under this head.

(b) The most important example of a true tensor product (and the example from which we chose the word "tensor") is probably the following. If G is the tangent vector space at a point of a differentiable manifold, then $G \circ G$ is the space of contravariant tensors of order 2 at the point. (Here $G \circ G$ is not the discrete, but the reduced, or topological, tensor product; see Part II or Part III. The same remark applies to other examples below.) Using also the "conjugate space" $L(G)$ and iterated tensor products gives tensors of all orders (see §11). Of course these spaces are usually defined in terms of coördinate systems in G.

Note that in terms of a fixed coördinate system, $G \circ G$ gives: vector times vector equals matrix. For a generalization, see (i) below.

(c) If G in (b) is replaced by Hilbert space, $G \circ G$ is a Hilbert space,[7] except for the completeness postulate (which could be obtained by completing the space or allowing certain infinite sums in $G \circ G$).

(d) The true tensor product $G \circ H$ has also been used in case one of G, H has a finite number of generators, and has been applied in topology.[8] From the examples (j) and Theorems 3 and 5 below, we may at once determine $G \circ H$ if both G and H have finite sets of generators.

The remaining examples are in general not true tensor products, but come under the heading (a). The general case $G \circ H \rightarrow Z$ does not often occur. The case $G \circ G \rightarrow Z$ appears in (b). The cases $G \circ H \rightarrow H$ and $G \circ G \rightarrow G$ appear in (e) and (g) below.

(e) If G is a group, with "operators" from the group R, i.e., $r \cdot g = g'$, the distributive laws are generally assumed; we have $R \circ G \rightarrow G$. Here one generally lets R be a ring (see §6).

(f) If G is a group and R is a ring, and we wish to form from G a group G^*

[7] See F. J. Murray and J. von Neumann, *On rings of operators*, Annals of Mathematics, vol. 37(1936), pp. 116–229, Chapter I. As a bounded operator A in G corresponds uniquely to an element f in G: $A(g) = (f, g)$, their space $G \otimes G$ corresponds to our $G \circ G$. M. H. Stone and J. W. Calkin have also considered a direct definition of $G \circ G$ such as we give. Compare also M. Kerner, *Abstract differential geometry*, Compositio Mathematica, vol. 4 (1937), pp. 308–341.

[8] See Alexandroff-Hopf, *Topologie* I, pp. 585–586 and p. 233, (15), and H. Freudenthal, Fundamenta Mathematicae, vol. 29(1937). The definition of $G \circ H$ is indirect. The case that one of G, H is a free group has been studied by H. Freudenthal, Compositio Mathematica, vol. 4(1937), pp. 145–234, Chapter III.

which "admits" R as operator ring, we need merely use $G^* = R \circ G$ (see Theorem 12 below). If we wish to replace G by a group G^* in which division by any integer $\neq 0$ is possible and unique, we use $G^* = Rt \circ G$ (see §8).

(g) If G is a group, any choice of $G \circ G \to G$ makes G a ring (in general non-associative), and conversely.

(h) Let V_p, V_q and V_r be linear spaces (= vector spaces) of dimensions p, q and r. Set $G = Ch_{V_q}(V_p)$ (= group of linear maps of V_p into V_q), $H = Ch_{V_r}(V_q)$, $Z = Ch_{V_r}(V_p)$. Obviously, we have $G \circ H \to Z$. G, H, Z, and $G \circ H$ are vector spaces of dimensions pq, qr, pr, and pq^2r. Hence $Z \approx G \circ H$ is possible only if $q = 1$, i.e., $V_q \approx Rl$. In this case it is true, as shown by (10.7) and (10.11) below. If we choose fixed coördinate systems in V_p, V_q and V_r, then G, H and Z may be interpreted as groups of matrices.

(i) If $G = H$ is the (additive) group of continuous functions $g(x)$, $0 \leqq x \leqq 1$, we may interpret $G \circ H$ as a subgroup of the group of continuous functions $z(x, y)$, $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$, with $g \cdot h$ corresponding to $z(x, y) = g(x)h(y)$. As is well known from the theory of integral equations, if we allow infinite sums, we may obtain all continuous functions $z(x, y)$.

(j) Finally, we give some examples of tensor products, using the groups most commonly used as coefficient groups in topology. Let Rt_1 and Rl_1 be Rt and Rl reduced mod 1.

$$I_0 \circ G \approx G, \qquad I_\mu \circ G \approx G_\mu \qquad \text{(Theorems 7, 8)},$$

$$I_\mu \circ I_\nu \approx I_{(\mu,\nu)},$$

$$I_\mu \circ Rt \approx I_\mu \circ Rl \approx I_\mu \circ Rt_1 \approx I_\mu \circ Rl_1 \approx 0 \qquad (\mu > 0),$$

$$Rt \circ Rt \approx Rt, \qquad Rt \circ Rl \approx Rl \circ Rl \approx Rl,$$

$$Rt \circ Rt_1 \approx Rt \circ Rl_1 \approx Rt_1 \circ Rt_1, \text{ etc.}, \approx 0.$$

5. **General properties.** We first consider commutative and associative properties.

THEOREM 3. *There is a natural isomorphism $G \circ H \approx H \circ G$, given by*

$$\phi(\textstyle\sum g_i \cdot h_i) = \sum h_i \cdot g_i. \tag{5.1}$$

THEOREM 4. *There are natural isomorphisms*

$$F \circ (G \circ H) \approx F \circ G \circ H \approx (F \circ G) \circ H,$$

where $F \circ G \circ H$ is the group of all $\sum f_i \cdot g_i \cdot h_i$, using the three distributive laws. The isomorphisms are given by

$$\phi(\textstyle\sum f_i \cdot g_i \cdot h_i) = \sum (f_i \cdot g_i) \cdot h_i, \qquad \psi(\sum f_i \cdot g_i \cdot h_i) = \sum f_i \cdot (g_i \cdot h_i). \tag{5.2}$$

The first theorem is evident; we prove the second, using ϕ. The definition of ϕ is unique, as any equivalence relation in the $\sum f_i \cdot g_i \cdot h_i$ corresponds to one in the $\sum (f_i \cdot g_i) \cdot h_i$. If $\phi(\sum f_i \cdot g_i \cdot h_i) = 0$, then an equivalence sequence carries

$\sum(f_i \cdot g_i) \cdot h_i$ into 0; a corresponding sequence carries $\sum f_i \cdot g_i \cdot h_i$ into 0; hence ϕ is an isomorphism into a subgroup of $(F \circ G) \circ H$. Finally, given any

$$\sum_i z_i \cdot h_i = \sum_i \left(\sum_j f_{ij} \cdot g_{ij}\right) \cdot h_i = \sum_{i,j} (f_{ij} \cdot g_{ij}) \cdot h_j$$

in $(F \circ G) \circ H$, ϕ carries $\sum f_{ij} \cdot g_{ij} \cdot h_i$ into it. This completes the proof.

Next we prove the distributive laws with respect to direct sums and difference groups.

THEOREM 5. *There is a natural isomorphism*

$$(F \oplus G) \circ H = F \circ H \oplus G \circ H,$$

given by

$$\begin{aligned}(5.3)\quad &\phi[(f_1, g_1) \cdot h_1 + \cdots + (f_n, g_n) \cdot h_n] \\ &\qquad = (f_1 \cdot h_1 + \cdots + f_n \cdot h_n, g_1 \cdot h_1 + \cdots + g_n \cdot h_n).\end{aligned}$$

To show that ϕ is uniquely defined, we have, for instance, as $(f, g) + (f', g') = (f + f', g + g')$,

$$\begin{aligned}\phi[\cdots + (f, g) \cdot h + (f', g') \cdot h + \cdots] \\ &= (\cdots + f \cdot h + f' \cdot h + \cdots, \cdots + g \cdot h + g' \cdot h + \cdots) \\ &= (\cdots + (f + f') \cdot h + \cdots, \cdots + (g + g') \cdot h + \cdots) \\ &= \phi[\cdots + \{(f, g) + (f', g')\} \cdot h + \cdots].\end{aligned}$$

ϕ maps the first group into the whole of the second; for

$$(5.4)\quad \phi[(f_1, 0) \cdot h_1 + \cdots + (0, g_1) \cdot h_1' + \cdots] = (f_1 \cdot h_1 + \cdots, g_1 \cdot h_1' + \cdots).$$

Clearly ϕ is a homomorphism. Now suppose $\phi(\alpha) = 0$; let α be given as in (5.3). First, we may transform α into the form of the left side of (5.4). For each half of the right side of (5.3), there is an equivalence sequence carrying it into 0. There are corresponding sequences acting on the left side of (5.4), which shows that $\alpha = 0$. Hence ϕ is an isomorphism.

THEOREM 6. *If G' is a subgroup of G, there is a natural isomorphism*

$$(G \ominus G') \circ H = G \circ H \ominus \sum\nolimits^* (G' \cdot H),$$

given as follows. If ψ and Ψ are the natural homomorphisms of G into $G \ominus G'$ and of $G \circ H$ into $G \circ H \ominus \sum^*(G' \cdot H)$, we set

$$(5.5)\quad \phi[\psi(g_1) \cdot h_1 + \cdots + \psi(g_n) \cdot h_n] = \Psi(g_1 \cdot h_1 + \cdots + g_n \cdot h_n).$$

By Theorem 3, there is a similar relation with G and H interchanged.

To show that ϕ is uniquely defined, suppose first that $\psi(g_1) = \psi(\bar{g}_1)$. Then $\bar{g}_1 = g_1 + g'$ (g' in G'), and

$$\Psi(\bar{g}_1 \cdot h_1 + \cdots) = \Psi(g_1 \cdot h_1 + \cdots) + \Psi(g' \cdot h_1) = \Psi(g_1 \cdot h_1 + \cdots).$$

The rest of the proof is like previous proofs. For instance, if the element (5.5) vanishes, then $\sum g_i \cdot h_i$ is in $\sum^*(G' \cdot H)$, and hence may be transformed into the form $\sum g_i' \cdot h_i'$ (g_i' in G'). The same transformations may be carried out on the left side of (5.5); as $\psi(g_i') = 0$, this gives $\sum \psi(g_i) \cdot h_i = 0$.

Remark. $\sum^*(G' \cdot H)$ is perhaps "smaller" than $G' \circ H$; for instance, if $G = I_0$, $G' = 2G$, $H = I_2$, then $G' \circ H \approx I_2$, $\sum^*(G' \cdot H) \approx 0$. But *there is a natural homomorphism of $G' \circ H$ onto the whole of $\sum^*(G' \cdot H)$*, clearly. Compare Theorem 28, Part II.

THEOREM 7. *There is a natural isomorphism $I_0 \circ G \approx G$, given by*

$$\phi\left(\sum a_i \cdot g_i\right) = \sum a_i g_i . \tag{5.6}$$

The proof is like previous proofs. Note that we have a normal form for elements of $I_0 \circ G$: if we use Theorem 2,

$$\sum a_i \cdot g_i = \sum 1 \cdot a_i g_i = 1 \cdot \sum a_i g_i = 1 \cdot g'. \tag{5.7}$$

The expression of an element in the normal form is unique, by the theorem.

THEOREM 8. *There is a natural isomorphism $I_\mu \circ G \approx G_\mu$, given by*[9]

$$\phi\left(\sum_i a_\mu^i \cdot g^i\right) = \sum_i a^i g_\mu^i. \tag{5.8}$$

Using Theorems 6 and 7, we see easily that the following isomorphism is the one given by the theorem:

$$I_\mu \circ G = (I_0 \ominus \mu I_0) \circ G \approx I_0 \circ G \ominus \sum{}^* (\mu I_0 \cdot G)$$
$$= I_0 \circ G \ominus \sum{}^* (I_0 \cdot \mu G) \approx I_0 \circ (G \ominus \mu G) \approx G_\mu .$$

THEOREM 9. *If G is completely divisible and every element of H is of finite order, then $G \circ H \approx 0$.*

For if $mh = 0$, then $g \cdot h = mg' \cdot h = g' \cdot mh = 0$.

THEOREM 10. *If G' and H' are subgroups of the nullifiers of H and G in G and H, respectively, then there are natural isomorphisms*

$$G \circ H \approx (G \ominus G') \circ H \approx G \circ (H' \ominus H) \approx (G \ominus G') \circ (H \ominus H');$$

if ϕ and ψ are the natural isomorphisms of G into $G \ominus G'$ and of H into $H \ominus H'$, these are given by

$$\sum g_i \cdot h_i \approx \sum \phi(g_i) \cdot h_i \approx \sum g_i \cdot \psi(h_i) \approx \sum \phi(g_i) \cdot \psi(h_i).$$

First, applying Theorem 6, we find, as $G' \cdot H = 0$,

$$G \circ H = G \circ H \ominus \sum{}^* (G' \cdot H) \approx (G \ominus G') \circ H, \text{ etc.}$$

Next, for any h' in H', $\phi(g) \cdot h'$ corresponds to $g \cdot h' = 0$ in the first isomorphism above; hence $(G \ominus G') \cdot H' = 0$, and

$$(G \ominus G') \circ H \approx (G \ominus G') \circ H \ominus \sum{}^* ((G \ominus G') \cdot H') \approx (G \ominus G') \circ (H \ominus H').$$

[9] g_μ is the element of G_μ corresponding to g in G.

We end by showing that the discrete tensor product of any two groups, not necessarily Abelian, is isomorphic to the discrete tensor product of the two groups "made Abelian".

THEOREM 11. *Let G and H be any two groups, and let G' and H' be their commutator subgroups. Then there is a natural isomorphism*

$$G \circ H \approx (G \ominus G') \circ (H \ominus H').$$

Because of Theorem 10, we need merely show that any commutator is in the nullifier of the other group; this follows at once from §3, (d).

6. **Sets, groups, rings, operators.** If A and B are two sets of elements, we may define their (discrete) tensor product as the set of all symbols $\pm a_1 \cdot b_1 \pm \cdots \pm a_n \cdot b_n$, with the obvious definition of $+$, which we assume commutative. This is a free group, generated by all $a \cdot b$; if A and B have m and n elements, respectively, then $A \circ B$ has mn generators.

If G is an Abelian group and A is a set of elements, their tensor product is the set of all $\sum g_i \cdot a_i$, with the distributive law as in (3.1), postulating that $+$ is commutative, and $0 \cdot a + g \cdot a' = g \cdot a'$. This is the "group of all linear forms over elements of A, with coefficients in G". An example is given by the groups of chains used in topology.

We shall say an Abelian group G *admits the ring R as operator ring*, or *admits* R simply, if R has a unit 1, and $rg = g'$ is defined satisfying

$$(6.1) \qquad \begin{aligned} r(g + g') &= rg + rg', \qquad (r + r')g = rg + r'g, \\ r(r'g) &= (rr')g \text{ or } (r'r)g, \qquad 1g = g. \end{aligned}$$

We call R a *left* or *right* operator according as we use $(rr')g$ or $(r'r)g$ in the third relation. In the second case, we might write gr in place of rg, obtaining $(gr')r = g(r'r)$. Suppose, for definiteness, we write $r[g]$ instead of rg. Then a ring can operate on itself in both ways, using

$$(6.2) \qquad r[r'] = rr' \quad \text{and} \quad r[r'] = r'r.$$

The associative law $r[r'[r'']] = (r[r'])[r'']$ holds in either case.

If G and H both admit R, to left or right, we say an isomorphism ϕ between G and H is an *operator isomorphism* if $\phi(rg) = r\phi(g)$; we use $\approx$ again, and say ϕ *preserves* the operator.

THEOREM 12. *If R is a ring with unit, and we define $R \circ G$, considering R as a group under addition, then $R \circ G$ admits R to left or right, under the definitions*

$$(6.3) \qquad r(\sum r_i \cdot g_i) = \sum rr_i \cdot g_i \quad \textit{or} \quad \sum r_i r \cdot g_i .$$

The proof is simple. The following theorem is a generalization.

THEOREM 13. *If G admits R to left or to right, then so does any tensor product $G \circ H$ or $H \circ G$, under the definition*

$$(6.4) \qquad r(\sum g_i \cdot h_i) = \sum rg_i \cdot h_i , \qquad r(\sum h_i \cdot g_i) = \sum h_i \cdot rg_i .$$

Suppose G and H both admit R, each to one side. Then we define the *reduced tensor product* $G \circ' H$ with respect to R as follows. Take the tensor product $G \circ H$, and define a new relation

$$rg \cdot h = g \cdot rh. \tag{6.5}$$

$G \circ' H$ is the group thus formed; it is the difference group of $G \circ H$ with the group generated by all $rg \cdot h - g \cdot rh$.

THEOREM 14. *If G admits R to the left, then there is a natural operator isomorphism*

$$R \circ' G \approx G,$$

letting R act on itself to the right and on $R \circ' G$ to the left, given by

$$\phi(\sum r_i \cdot g_i) = \sum r_i g_i . \tag{6.6}$$

Here, (6.5) is replaced by

$$rr' \cdot g = r'[r] \cdot g = r \cdot r'[g] = r \cdot r'g. \tag{6.5'}$$

To show that ϕ is uniquely defined, we have for instance

$$\phi(rr' \cdot g) = (rr')g = r(r'g) = \phi(r \cdot r'g).$$

ϕ is a homomorphism into the whole of G; for $\phi(1 \cdot g) = 1g = g$. It preserves the operator, for

$$\begin{aligned}\phi(r(\sum r_i \cdot g_i)) = \phi(\sum rr_i \cdot g_i) = \sum (rr_i)g_i &= \sum r(r_i g_i)\\ &= r(\sum r_i g_i) = r\phi(\sum r_i \cdot g_i).\end{aligned}$$

Finally, ϕ is (1-1). For if $\phi(\sum r_i \cdot g_i) = \sum r_i g_i = 0$, then

$$\sum r_i \cdot g_i = \sum 1 \cdot r_i g_i = 1 \cdot \sum r_i g_i = 1 \cdot 0 = 0.$$

The theorem clearly holds with "right" and "left" interchanged.

Suppose R and S are rings.[10] Then we can make $R \circ S$ a ring in four different ways, namely,

$$\begin{aligned}&(r \cdot s)(r' \cdot s') = rr' \cdot ss' \quad \text{or} \quad rr' \cdot s's, \text{ etc.},\\ &(\sum r_i \cdot s_i)(\sum r_j' \cdot s_j') = \sum \sum (r_i \cdot s_i)(r_j' \cdot s_j').\end{aligned} \tag{6.7}$$

The uniqueness of the definition is easily established. The associative and distributive laws hold. If R and S have units 1_R and 1_S, then so has $R \circ S$, namely, $1_R \cdot 1_S$.

We shall not discuss the questions of zero-divisors or of fields.

7. Rational multipliers and tensor products.

Definition. For any rational number r, $r = a/b$, $(a, b) = 1$, and any $A \subset G$

[10] Compare J. L. Dorroh, loc. cit.

(including $A = g$), we let rA be the set of all elements g' such that $bg' = ag$, g in A. This agrees with the definition of aA and with the natural definition of $(1/a)A$. Then some of the formal properties of rational numbers as multipliers hold. In particular, some elements can be divided by certain integers. *Division by integers, when it exists, is unique if and only if G has no elements* $\neq 0$ *of finite order.* For if g' and g'' are in $(1/a)g$, $g' \neq g''$, then $a(g' - g'') = g - g = 0$, so that $g' - g''$ is of finite order; if $g \neq 0$ is of finite order a, then $(1/a)0$ is not unique. We shall say G has *unique division* if it is completely divisible and has no elements $\neq 0$ of finite order. Because of Theorem 15 below, we may then multiply by rational numbers in such a group, and all formal laws will hold.

The only theorem we will need in §8 is the following.

THEOREM 15. *The following three statements are equivalent:*

(a) *G admits Rt as operator ring; we shall write* $r[g]$.

(b) *G has unique division.*

(c) *For each rational r and each g in G, rg is a unique element of G.*

Further, G can admit Rt in at most one way; if it does, then $rg = r[g]$.

First, if G admits Rt, then G has no elements of finite order. For, note first that (for $a > 0$, and hence for $a \leqq 0$),

$$(*) \qquad a[g] = (1 + \cdots + 1)[g] = 1[g] + \cdots + 1[g] = ag.$$

Now if $ag = 0$, $a \neq 0$, then $a[g] = ag = 0 = a0 = a[0]$; hence

$$g = 1[g] = \left(\frac{1}{a}\, a\right)[g] = \frac{1}{a}[a[g]] = \frac{1}{a}[a[0]] = 1[0] = 0.$$

Next, if (a) holds, then for each integer $a \neq 0$ and each g in G, $g' = (1/a)[g]$ exists, and $ag' = a[g'] = g$; hence (b) holds. (b) clearly implies (c). If (c) holds, then setting $r[g] = rg$ gives (a).

Finally, if two operations $r[g]$ and $r\{g\}$ are defined, then they agree; for by (*),

$$b\left(\frac{a}{b}\,[g]\right) = \left(b\,\frac{a}{b}\right)[g] = a[g] = ag = b\left(\frac{a}{b}\,\{g\}\right);$$

as G can have no elements of finite order, $(a/b)[g] = (a/b)\{g\}$. Also

$$b\left(\frac{a}{b}\,[g]\right) = a[g] = ag = b\left(\frac{a}{b}\,g\right),$$

and hence $r[g] = rg$.

Before considering tensor products, we consider some divisibility properties in general groups. Let δ_r denote the denominator of r; $\delta_r = b$ if $r = a/b$, $(a, b) = 1$.

LEMMA 1. *If rg is not void, then* $\delta_r \mid g$, *and conversely.*

For if $r = a/b$, $bg' = ag$, and $pa + qb = 1$, then

$$b(qg + pg') = qbg + pag = g.$$

The converse is clear.

LEMMA 2. *If* $(a, b) = 1$, *then*

$$\frac{a}{b} A = a\left(\frac{1}{b} A\right) = \frac{1}{b} (aA). \tag{7.1}$$

To prove the first relation, the elements of $a((1/b)A)$ are all g', $g' = ag^*$, g^* in $(1/b)A$, i.e., $bg^* = g$ in A; then $bg' = ag$, and as $(a, b) = 1$, g' is in $(a/b)A$. Conversely, if g' is in $(a/b)A$, then $bg' = ag$ (g in A). Choose p, q so that $pa + qb = 1$, and set $g^* = qg + pg'$. Then

$$bg^* = qbg + pag = g, \qquad ag^* = qbg' + pag' = g',$$

so that g^* is in $(1/b)A$ and g' is in $ag^* \subset a((1/b)A)$. The second relation is clear.

LEMMA 3. *For any integers a and b,*

$$\frac{1}{a}\left(\frac{1}{b} A\right) = \frac{1}{ab} A, \qquad a\left(\frac{1}{a} A\right) \subset A, \qquad \frac{1}{a} (aA) \supset A. \tag{7.2}$$

The proof is simple.

We turn now to tensor products.

LEMMA 4. *If* $\delta_r \mid g$ *and* $\delta_r \mid h$, *then*

$$g' \cdot h = g \cdot h' \quad \text{for any } g' \text{ in } rg \text{ and any } h' \text{ in } rh. \tag{7.3}$$

Set $r = a/b$, $(a, b) = 1$. If

$$bg' = ag, \qquad g = bg^*, \qquad bh' = ah, \qquad h = bh^*,$$

then

$$g \cdot h' = bg^* \cdot h' = g^* \cdot bh' = g^* \cdot ah = g^* \cdot abh^* = abg^* \cdot h^* = ag \cdot h^*$$
$$= bg' \cdot h^* = g' \cdot bh^* = g' \cdot h.$$

Example. If $\delta_r \mid h$ is false, $rg \cdot h$ may not be uniquely defined. For if $G = H = I_2$, $g = 0_2$, $h = 1_2$, then $G \circ H \approx I_2$, and $\frac{1}{2}g \cdot h$ contains both 0_2 and 1_2.

THEOREM 16. *If* $\delta_r \mid A$ *and* $\delta_r \mid B$, *then*

$$rA \cdot B = A \cdot rB; \tag{7.4}$$

if A and B are single elements, so is $rA \cdot B$.

This follows from Lemmas 1 and 4.

Remark. $r(g \cdot h)$ may be $\neq rg \cdot h$. For example, if $G = H = I_2$, $g = h = 0_2$, $r = \frac{1}{2}$, then $rg \cdot h = 0_2$, while $r(g \cdot h)$ contains both 0_2 and 1_2. However,

$$r(A \cdot B) \supset rA \cdot B; \tag{7.5}$$

for if $r = a/b$, $(a, b) = 1$, g in A, h in B, $bg' = ag$, so that $g' \cdot h$ is in $rA \cdot B$, then

$$b(g' \cdot h) = bg' \cdot h = ag \cdot h = a(g \cdot h) \text{ is in } a(A \cdot B),$$

so that $g' \cdot h$ is in $r(A \cdot B)$.

LEMMA 5. *If $b \mid A$ and $b \mid B$, then*

$$\frac{1}{b}(aA)\cdot B = \frac{a}{b}A\cdot B = a\left(\frac{1}{b}A\right)\cdot B = A\cdot\frac{1}{b}(aB) = A\cdot\frac{a}{b}B = A\cdot a\left(\frac{1}{b}B\right); \tag{7.6}$$

if A and B are single elements, so is the above.

Say $(a, b) = k$, $a = a'k$, $b = b'k$; then $(a', b') = 1$. To prove the first relation, we use Lemmas 2 and 3 and Theorem 16, and the fact $b \mid aA$:

$$\frac{1}{b}(aA)\cdot B = \frac{1}{b'}\left(\frac{1}{k}(k(a'A))\right)\cdot B \supset \frac{1}{b'}(a'A)\cdot B = \frac{a'}{b'}A\cdot B = \frac{a}{b}A\cdot B,$$

$$\frac{1}{b}(aA)\cdot B = A\cdot a\left(\frac{1}{b}B\right) = A\cdot a'\left(k\left(\frac{1}{k}\left(\frac{1}{b'}B\right)\right)\right) \subset A\cdot a'\left(\frac{1}{b'}B\right) = \frac{a}{b}A\cdot B.$$

From these the relation follows. The other relations are consequences of this one or are easily proved. The last statement follows from Theorem 16.

THEOREM 17. *If $\delta_r\delta_{r'} \mid A$ and $\delta_r\delta_{r'} \mid B$,[11] then*

$$r(r'A)\cdot B = (rr')A\cdot B = A\cdot(rr')B, \text{ etc.}; \tag{7.7}$$

if A and B are single elements, so is the above.

Say $r = a/b$, $r' = c/d$, $(a, b) = (c, d) = 1$. As $bd \mid cA$, etc.,

$$r(r'A)\cdot B = a\left(\frac{1}{b}\left(\frac{1}{d}(cA)\right)\right)\cdot B = \frac{1}{bd}(cA)\cdot aB = ac\left(\frac{1}{bd}A\right)\cdot B$$

$$= \frac{ac}{bd}A\cdot B = (rr')A\cdot B, \text{ etc.}$$

8. **The tensor product $Rt \circ G$.** First note that, if F is any completely divisible group (in particular, Rt), then in studying $F \circ G$, we could assume that G has no elements $\neq 0$ of finite order. For otherwise, let G' be the subgroup of elements of finite order of G. As G' is in the nullifier of F, $\sum^*(F\cdot G') = 0$ (see Theorem 9); hence, by Theorem 10,

$$F \circ G \approx F \circ (G \ominus G').$$

Thus we may replace G by $G \ominus G'$, which has no elements $\neq 0$ of finite order.

THEOREM 18. *In $Rt \circ G$, each element may be written in the form $(1/a)\cdot g$. If G has no elements $\neq 0$ of finite order, then $r\cdot g = 0$ if and only if $r = 0$ or $g = 0$.*

First,

$$\sum r_i\cdot g_i = \sum \frac{a_i}{a}\cdot g_i = \frac{1}{a}\cdot\sum a_i g_i = \frac{1}{a}\cdot g.$$

Next, suppose we have an equivalence sequence reducing $r\cdot g$ to $0\cdot 0$. In all terms occurring, there is a least common denominator c. Multiplying every-

[11] Possibly this hypothesis can be weakened.

thing by c gives an equivalence sequence, which may be interpreted as a sequence in $I_0 \circ G$, or again, in G itself. Hence, if $r = a/b$, we have $(ca/b)g = 0$. If $r \neq 0$, then $ca/b \neq 0$, and as G has no elements of finite order, $g = 0$.

THEOREM 19. *$Rt \circ G$ has unique division.*

This follows from Theorems 12 and 15.

THEOREM 20. *There is an isomorphism $G \approx Rt \circ G$, given by $\phi(\sum r_i \cdot g_i) = \sum r_i g_i$, if and only if G has unique division.*

This is an extension of Theorem 15. One half follows from Theorem 19; the other half is clear.

THEOREM 21. *If G has no elements $\neq 0$ of finite order, then $Rt \circ G$ is the smallest completely divisible group containing G. That is, if H is completely divisible and contains a subgroup $H_1 \approx G$, then it contains a subgroup $H_2 \approx Rt \circ G$.*

Let H' be the subgroup of elements of finite order of H. Clearly H' is completely divisible; hence we may write $H = H' \oplus H''$.[12] For any $h = h' + h''$, write $h' = \phi(h)$, $h'' = \psi(h)$; then ϕ and ψ are homomorphisms. Set $H_1'' = \psi(H_1)$; then $H_1'' \approx G$. For if $\psi(h_1) = 0$ (h_1 in H_1), then h_1 is in H', and hence is of finite order; but h_1 is in $H_1 \approx G$, which gives $h_1 = 0$.

Let H_2 be the subgroup of H'' containing all elements with multiples in H_1''. H_2 is completely divisible. For given h in H_2 and an integer $a \neq 0$, choose h^* in H so that $ah^* = h$, and set $h_1 = \psi(h^*)$. Then h_1 is in H'', and as h is in H'',

$$ah_1 = a\psi(h^*) = \psi(ah^*) = \psi(h) = h;$$

hence h_1 is in H_2. As H'' has no elements $\neq 0$ of finite order, neither has H_2; hence H_2 has unique division.

Let θ be the isomorphism of G into H_1''. As rh is uniquely defined for h in the group H_2 (Theorem 15), and clearly obeys $(r + r')/h = rh + r'h$, $r(h + h') = rh + rh'$, we may set

$$\Theta(\sum r_i \cdot g_i) = \sum r_i \theta(g_i),$$

defining a homomorphism of $Rt \circ G$ into H_2. Suppose $\Theta(\alpha) = 0$. If $\alpha = (1/a) \cdot g$ (Theorem 18), then $\Theta(\alpha) = (1/a)\theta(g) = 0$. Multiplying by a gives $\theta(g) = 0$, and hence $g = 0$, and $\alpha = 0$, as θ is an isomorphism. Hence Θ is (1-1). For any h in H_2, we may take a so that ah is in H_1''; then for some g, $ah = \theta(g) = \Theta(1 \cdot g)$, and $h = \Theta((1/a) \cdot g)$; hence Θ is an isomorphism, and the theorem is proved.

9. **Tensor products and character groups.** In some cases, the group $Ch_H(G)$ of homomorphisms of G into H can be expressed in terms of the two groups H and $Ch_{I_0}(G)$, by (9.1). See also Theorem 25 of Part II. We remark in passing that $Ch_H(G)$ and G form a group pair with respect to H, with the definition $\Phi(\sum \phi_i \cdot g_i) = \sum \phi_i(g_i)$ (ϕ_i in $Ch_H(G)$, g_i in G).

[12] See R. Baer, *The subgroup of elements of finite order of an Abelian group*, Annals of Mathematics, vol. 37(1936), pp. 766–781, (1; 1).

THEOREM 22.[13] *There is a natural isomorphism*

$$Ch_{I_0}(G) \circ H \approx Z \subset Ch_H(G), \tag{9.1}$$

defined as follows. For u_i in $Ch_{I_0}(G)$ and h_i in H,

$$\Phi(\textstyle\sum u_i \cdot h_i\,;\, g) = \sum u_i(g) h_i\,. \tag{9.2}$$

If either G or H is a free group with a finite number of generators, then $Z = Ch_H(G)$.

It is clear that the definition of Φ is unique, and Φ is a homomorphism. We must show that it is (1-1). Suppose the element (9.2) equals 0. Say the sum contains n terms. Let $A = I_0 \oplus \cdots \oplus I_0$ be the group of all n-tuples $(a_1, \cdots, a_n)$ of integers, and let A' be the subgroup of all $(a_1, \cdots, a_n)$ in A for which $\sum a_i h_i = 0$. We may choose a base

$$\alpha_1, \cdots, \alpha_n\,; \qquad \alpha_i = (a_{i1}, \cdots, a_{in})$$

in A and integers $p_1, \cdots, p_m$ $(m \leqq n)$ such that

$$p_1\alpha_1, \cdots, p_m\alpha_m$$

form a base in A'.[14] For each g, let $u(g)$ be the element $(u_1(g), \cdots, u_n(g))$ of A; as $\sum u_i(g)h_i = 0$, $u(g)$ is in A'. Hence, for each g, there is a uniquely defined set of numbers $\rho_1(g), \cdots, \rho_m(g)$ such that

$$u(g) = \sum_{j=1}^{m} \rho_j(g) p_j \alpha_j\,;$$

hence

$$u_i = \sum_{j=1}^{m} p_j a_{ji} \rho_j\,.$$

As the $u_i(g)$ are homomorphisms, so are $u(g)$ and the $\rho_i(g)$; the $\rho_i(g)$ are in $Ch_{I_0}(G)$. Set

$$\bar{h}_i = \sum_{k=1}^{n} a_{ik} h_k \qquad (i = 1, \cdots, m);$$

then

$$p_i \bar{h}_i = \sum_{k=1}^{n} p_i a_{ik} h_k = 0 \qquad (i = 1, \cdots, m),$$

by the choice of the α_i and p_i. Hence, using the distributive laws in $Ch_{I_0}(G) \circ H$,

$$\sum_{i=1}^{n} u_i \cdot h_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} p_j a_{ji} \rho_j \right) \cdot h_i = \sum_{j=1}^{m} \left(\rho_j \cdot p_j \sum_{i=1}^{n} a_{ji} h_i \right)$$
$$= \sum_{j=1}^{m} (\rho_j \cdot p_j \bar{h}_j) = \sum_{j=1}^{m} (\rho_j \cdot 0) = 0,$$

as required.

[13] Compare Theorem 25.

[14] See, for example, Alexandroff-Hopf, loc. cit., p. 566.

Now suppose H has a base $\bar{h}_1, \cdots, \bar{h}_n$, so any h may be written uniquely $\sum a_i \bar{h}_i$. Let ϕ be any homomorphism of G into H; then we may write

$$\phi(g) = \sum u_i(g)\bar{h}_i,$$

and the $u_i(g)$ are elements of $Ch_{I_0}(G)$. Also

$$\Phi(\sum u_i \cdot \bar{h}_i; g) = \sum u_i(g)\bar{h}_i = \phi(g),$$

so Φ maps $Ch_{I_0}(G) \circ H$ into the whole of $Ch_H(G)$.

Suppose finally that G has a base $\bar{g}_1, \cdots, \bar{g}_n$. Let $\bar{u}_i(g)$ be the element of $Ch_{I_0}(G)$ defined by $\bar{u}_i(\bar{g}_i) = 1$, $\bar{u}_i(\bar{g}_j) = 0$ $(j \neq i)$. Take any homomorphism ϕ of G into H. Then for any $g = \sum a_i \bar{g}_i$, $\bar{u}_i(g) = a_i$, and

$$\phi(g) = \sum a_i \phi(\bar{g}_i) = \sum \bar{u}_i(g)\phi(\bar{g}_i);$$

hence, setting $h_i = \phi(\bar{g}_i)$,

$$\Phi(\sum \bar{u}_i \cdot h_i; g) = \sum \bar{u}_i(g)\phi(\bar{g}_i) = \phi(g).$$

This completes the proof.

Examples. Suppose $G = H = I_2$. Then $Ch_H(G)$ has two elements, while $Ch_{I_0}(G) \circ H$ has only one. Again, let G be the additive group of triadic rational numbers (all numbers of the form $a/3^b$), and set $H = I_2$. There are two elements in $Ch_H(G)$, determined by $\phi(1) = 0_2$ and $\phi(1) = 1_2$; but there is only one element in $Ch_{I_0}(G) \circ H$.

II. Linear spaces

10. **Products, finite dimensional spaces.** A *linear space*, or *vector space*, G, is an Abelian[15] group which admits the real numbers Rl as operators (see §6). Let $G(g_1, \cdots, g_m)$ be the subspace of G generated by $g_1, \cdots, g_m$, i.e., all $\sum a_i g_i$ (a_i real). If such a set generates G itself, then let $g_1, \cdots, g_m$ be such a set with the least number of elements. Then these elements form a *base* for G, and G is of *dimension* m.

In any finite dimensional linear space G, with a base $g_1, \cdots, g_m$, we may introduce a *natural topology* by defining neighborhoods $U(\epsilon)$ of 0 for each $\epsilon > 0$, consisting of all $\sum a_i g_i$ with $\sum a_i^2 < \epsilon^2$. The topology is independent of the choice of a base. In this topology, the operation ag is continuous in both variables.

In the tensor product $G \circ H$, we clearly wish to have

$$a(g \cdot h) = ag \cdot h = g \cdot ah \qquad (a \text{ in } Rl); \tag{10.1}$$

hence we use the *reduced tensor product* (see (6.4)), but call it the *tensor product* simply. Without this, we would have for instance in Rl, $\sqrt{2} \cdot 1 \neq 1 \cdot \sqrt{2}$. Further, *if we assume that* $g \cdot h$ *is continuous, then* (10.1) *follows.* To show this, the last statement in Theorem 15, and Theorem 16, show that $bg \cdot h = g \cdot bh$ for any rational b. Letting $b \to a$ gives the result.

[15] The group is necessarily Abelian. Compare §3, (e). If G is not linear, it can be made so by taking $Rl \circ G$; see Theorem 12, §6.

We assume in the rest of §10 that G and H have bases $\bar{g}_1, \cdots, \bar{g}_m$ and $\bar{h}_1, \cdots, \bar{h}_n$, respectively.

THEOREM 23. *An element of $G \circ H$ may be written uniquely in any one of the three normal forms*

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(\bar{g}_i\cdot\bar{h}_j) = \sum_{i=1}^{m} \bar{g}_i\cdot h_i' = \sum_{j=1}^{n} g_j'\cdot\bar{h}_j. \tag{10.2}$$

For, if

$$g_k = \sum b_{ki}\bar{g}_i, \qquad h_k = \sum c_{ki}\bar{h}_i, \tag{10.3}$$

then the distributive laws give

$$\sum_k g_k\cdot h_k = \sum_k \left(\sum_i b_{ki}\bar{g}_i\right)\cdot\left(\sum_j c_{kj}\bar{h}_j\right) = \sum_k\sum_i \left(\bar{g}_i\cdot\sum_j b_{ki}c_{kj}\bar{h}_j\right)$$

$$= \sum_i \bar{g}_i\cdot\sum_j\sum_k b_{ki}c_{kj}\bar{h}_j, \text{ etc.};$$

thus (10.2) holds with

$$a_{ij} = \sum_k b_{ki}c_{kj}, \qquad h_i' = \sum_j a_{ij}\bar{h}_j, \qquad g_i' = \sum_j a_{ji}\bar{g}_j. \tag{10.4}$$

Given any expression $\sum g_i\cdot h_i$ for α in $G \circ H$, the above procedure gives the normal forms in a unique manner; we must show that if $\sum g_i\cdot h_i = \sum g_i^*\cdot h_i^*$, the two expressions give the same result. It is sufficient to prove this for $(g + g^*)\cdot h$ and $g\cdot h + g^*\cdot h$, for $g\cdot(h + h^*)$ and $g\cdot h + g\cdot h^*$, and for $ag\cdot h$ and $g\cdot ah$. In each case, the proof is simple.

Let $Ch_H(G)$ denote the group of linear maps (= continuous homomorphisms) of G into H; this is a linear space of dimension mn. In particular, $L(G) = Ch_{Rl}(G)$ is the group of linear real-valued functions in G, and is called the *conjugate space* of G. Here, isomorphism will mean continuous isomorphism = operator isomorphism. The following theorem is well known.

THEOREM 24. $L(G) = G$. *Further, there is a natural isomorphism*

$$L(L(G)) \approx G, \tag{10.5}$$

defined as follows. For any g in G, $\phi(g)$ is the element of $L(L(G))$ which, for any u in $L(G)$, has the value $u(g)$.

Let $\bar{u}_i(g)$ be the element of $L(G)$ such that $\bar{u}_i(\bar{g}_i) = 1$, $\bar{u}_i(\bar{g}_j) = 0$ $(j \neq i)$. Clearly $\bar{u}_1, \cdots, \bar{u}_m$ form a base in $L(G)$; hence $L(G) \approx G$. Next, ϕ is linear. It is (1-1); for if $\phi(g) = 0$, then $u(g) = 0$ (all u in $L(G)$), which implies $g = 0$. Given any v in $L(L(G))$, set $a_i = v(\bar{u}_i)$; then for any $u = \sum b_i\bar{u}_i$,

$$u\left(\sum a_i\bar{g}_i\right) = \sum_i a_i \sum_j b_j\bar{u}_j(\bar{g}_i) = \sum a_ib_i = \sum b_iv(\bar{u}_i) = v(u),$$

so that $\phi(\sum a_i\bar{g}_i) = v$. Clearly $\phi(ag) = a\phi(g)$; hence ϕ is an isomorphism.

THEOREM 25.[16] *There is a natural isomorphism*

$$Ch_H(G) \approx L(G) \circ H, \tag{10.6}$$

given by

$$\phi(\sum u_i \cdot h_i ; g) = \sum u_i(g) h_i . \tag{10.7}$$

ϕ is clearly uniquely defined. If we write all elements of $L(G) \circ H$ in the third normal form $\sum u_i \cdot \bar{h}_i$, the properties of ϕ are easily established; for any element of $Ch_H(G)$ can be written uniquely as $\sum u_i(g)\bar{h}_i$, and if this is the zero element, i.e., it is equal to zero in H for all g, then all $u_i(g) = 0$.

COROLLARY I. *$G \circ H$ may be written in the form*

$$G \circ H \approx L(L(G)) \circ H \approx Ch_H(L(G)). \tag{10.8}$$

The isomorphism of the first group into the last is given as follows. For $\sum g_i \cdot h_i$ in $G \circ H$ and u in $L(G)$,

$$\phi(\sum g_i \cdot h_i ; u) = \sum u(g_i) h_i . \tag{10.9}$$

COROLLARY II. *There is a natural isomorphism*

$$Ch_G(Rl) \approx G; \tag{10.10}$$

for u in $Ch_G(Rl)$, $\phi(u) = u(1)$.

For $L(Rl) \circ G \approx Rl \circ G \approx G$. (Moreover, a direct proof is obvious.)

THEOREM 26. *$G \circ H$ is a linear space of dimension mn, with a base $\bar{g}_1 \cdot \bar{h}_1, \cdots, \bar{g}_m \cdot \bar{h}_n$. If $\{U\}$ and $\{V\}$ are neighborhood systems in G and H, respectively, defining their natural topologies, then either of the following neighborhood systems, if we use $p = \min(m, n)$,*

$$N(U, V) = U \cdot V + \cdots + U \cdot V \qquad (p \text{ summands}), \tag{10.11}$$

$$N(U_1, U_2, \cdots ; V_1, V_2, \cdots) = \sum_k{}^*(U_k \cdot V_k) \tag{10.12}$$

defines the natural topology in $G \circ H$.[17] *The multiplication $g \cdot h$ is continuous.*

The first part of the theorem follows from Theorem 23. Let N, N', N'' denote natural neighborhoods and those of (10.11) and (10.12). Given an $N = N(\epsilon)$, consisting of all $\sum a_{ij} \bar{g}_i \cdot \bar{h}_j$ with $\sum a_{ij}^2 \leqq \epsilon^2$, set $\epsilon_1 = \epsilon/(mn)^{\frac{1}{2}}$, and let

$$U_k = U(\epsilon_1/2^k), \qquad V_k = V(1), \qquad (k = 1, 2, \cdots),$$

be natural neighborhoods in G and H. Then if $g_k = \sum b_{ki}\bar{g}_i$ is in U_k and $h_k = \sum c_{kj}\bar{h}_j$ is in V_k, (10.4) gives, if we use any finite number ν of summands in (10.12),

$$|a_{ij}| = \left| \sum_k b_{ki} c_{kj} \right| < \sum_{k=1}^{\nu} \epsilon_1/2^k < \epsilon_1 = \epsilon/(mn)^{\frac{1}{2}}.$$

[16] This holds if at least one of G, H is of finite dimension. Compare Theorem 22.

[17] If we map Rl into a curve everywhere dense on the torus, the topology of the torus gives an "unnatural" topology in Rl. In $Rl \circ Rl$, either type of neighborhood as here given then contains the whole space.

Hence $\sum a_{ij}^2 < \epsilon^2$, and $\sum_{k=1}^{p} g_k \cdot h_k$ is in N. Thus any N contains an N''.

Next, given an N'', take

$$U \subset U_1 \cap \cdots \cap U_p, \qquad V \subset V_1 \cap \cdots \cap V_p.$$

Then clearly $N' = N(U, V) \subset N''$.

Next, take any $N' = N(U, V)$. Suppose for definiteness that $p = m$. Take ϵ_1 so that $U(2\epsilon_1) \subset U$ and $V(\epsilon_1) \subset V$, and set $\epsilon = \epsilon_1^2$. Now take any α of $G \circ H$ in $N(\epsilon)$; then we can write $\alpha = \sum a_{ij} \bar{g}_i \cdot \bar{h}_j$, with $\sum a_{ij}^2 < \epsilon^2$. Also,

$$\alpha = \sum_{i=1}^{m} \left(\epsilon_1 \bar{g}_i \cdot \sum_{j=1}^{n} \theta a_{ij} \bar{h}_j \right), \qquad \theta = \frac{1}{\epsilon_1}.$$

As $m = p$ and $\epsilon_1 \bar{g}_i$ is in $U(2\epsilon_1) \subset U$, to show that $N(\epsilon) \subset N(U, V)$, it is sufficient to show that $\sum_j \theta a_{ij} \bar{h}_j$ is in $V(\epsilon_1)$. But

$$\sum_j \theta^2 a_{ij}^2 \leqq \theta^2 \sum_{i,j} a_{ij}^2 < \theta^2 \epsilon^2 = \epsilon_1^2,$$

and this proves the statement.

The continuity of $g \cdot h$ is clear from the relation

$$\sum (a_i + a_i') \bar{g}_i \cdot \sum (b_j + b_j') \bar{h}_j - \sum a_i \bar{g}_i \cdot \sum b_j \bar{h}_j = \sum (a_i' b_j + a_i b_j' + a_i' b_j') \bar{g}_i \cdot \bar{h}_j.$$

If G and H are metric, and hence scalar products $g \circ g'$ and $h \circ h'$ are defined, we may define scalar products and hence a metric in $G \circ H$ by

(10.13) $$\left(\sum_k g_k \cdot h_k\right) \circ \left(\sum_l g_l' \cdot h_l'\right) = \sum_{k,l} (g_k \circ g_l')(h_k \circ h_l').$$ [18]

11. **Tensor algebra.** Let G be a linear space of finite dimension n; in §12, it will be the "tangent space" at a point of a manifold. Any element of G we shall call a *contravariant vector*. An element of $H = L(G)$ we call a *covariant vector*. Any element of the linear space

(11.1) $$T(p, q) = G \circ \cdots \circ G \circ H \circ \cdots \circ H \qquad (p \text{ factors } G, q \text{ factors } H)$$

we shall call a *tensor of contravariant order p and covariant order q*. As $L(p, q)$ is a linear space, we may *add* two tensors of the same type, and *multiply* a tensor by a real number. Using Theorems 3 and 4 in Part I, we have

$$(G \circ \cdots \circ H \circ \cdots) \circ (G \circ \cdots \circ H \circ \cdots)$$
$$= G \circ \cdots \circ G \circ \cdots \circ H \circ \cdots \circ H \circ \cdots.$$

Hence a tensor of $T(p, q)$ and a tensor of $T(p', q')$ may be *multiplied*, giving a tensor of $T(p + p', q + q')$.

The process of *contraction* is as follows. To contract the element $g \cdot h$ of

[18] For a study of this metric in Hilbert spaces, see Murray and von Neumann, loc. cit.

$T(1, 1) = G \circ H$, recall that $H = L(G)$, and set $\phi(g \cdot h) = h(g)$, a real number. To contract the element

$$\alpha = \sum_k g_k^1 \cdots g_k^p \cdot h_k^1 \cdots h_k^q \quad \text{of } T(p, q)$$

with respect to the p-th g and the q-th h, for example, set

$$\phi(\alpha) = \sum_k h_k^q(g_k^p) g_k^1 \cdots g_k^{p-1} \cdot h_k^1 \cdots h_k^{q-1}; \tag{11.2}$$

this is an element of $T(p - 1, q - 1)$.

Let $\bar{g}_1, \cdots, \bar{g}_n$ form a base in G, and choose $\bar{h}^i$ so that $\bar{h}^i(\bar{g}_j) = \delta_j^i$; then $\bar{h}^1, \cdots, \bar{h}^n$ form a base in H. By the proof of Theorem 23, we may write any element of $T(p, q)$ uniquely in the normal form

$$\alpha = \sum_{i_r, j_s = 1}^{n} A_{j_1 \cdots j_q}^{i_1 \cdots i_p} \bar{g}_{i_1} \cdots \bar{g}_{i_p} \bar{h}^{j_1} \cdots \bar{h}^{j_q}; \tag{11.3}$$

there are n^{p+q} terms in the sum, and the $A_{j_1 \cdots j_q}^{i_1 \cdots i_p}$ are called the *components of* α in the coördinate system of the $\bar{g}_i$. Let us verify the *laws of transformation* of the components. Suppose we introduce the new base $g_1', \cdots, g_n'$. Say

$$\bar{g}_i = \sum_{k=1}^{n} a_i^k g_k', \qquad g_i' = \sum_{k=1}^{n} a_i'^k \bar{g}_k.$$

If $h'^i(g_j') = \delta_j^i$, then setting $h'^i = \sum b_k^i \bar{h}^k$ gives

$$\delta_j^i = h'^i(g_j') = \sum_k b_k^i \bar{h}^k \left(\sum_l a_j'^l \bar{g}_l\right) = \sum_{k,l} b_k^i a_j'^l \delta_l^k = \sum_k b_k^i a_j'^k.$$

Hence $b_j^i = a_j^i$, and $\bar{h}^i = \sum a_k'^i h'^k$. Putting in (11.3) and using the distributive laws gives

$$\alpha = \sum A_{j_1 \cdots j_q}^{i_1 \cdots i_p} a_{i_1}^{k_1} \cdots a_{i_p}^{k_p} a_{l_1}'^{j_1} \cdots a_{l_q}'^{j_q} g_{k_1}' \cdots g_{k_p}' \cdot h'^{l_1} \cdots h'^{l_q}.$$

Calling the new components $A_{l_1 \cdots l_q}'^{k_1 \cdots k_p}$, we have the ordinary laws of transformation. Note that

$$h(g) = \sum_i B_i \bar{h}^i \left(\sum_j A^j \bar{g}_j\right) = \sum_{i,j} A^j B_i \bar{h}^i(\bar{g}_j) = \sum_i A^i B_i,$$

so that the terms as here introduced agree with the usage in tensor algebra.

12. **Tensor analysis.** Let M be a differentiable manifold.[19] By a *parametrized curve* C starting at the point x_0 in M we shall mean a differentiable map ϕ of an interval $0 \leqq t \leqq \eta$ into M, with $\phi(0) = x_0$. Let us introduce a coördinate system into a neighborhood of M about x_0, i.e., a (1-1) differentiable map θ of a region of the space E of sets of n numbers $(x^1, \cdots, x^n)$ into M, with non-vanishing Jacobian; say $\theta(0, \cdots, 0) = x_0$. Then C translates into a curve C' in E,

[19] See, for instance, O. Veblen-J. H. C. Whitehead, *Foundations of Differential Geometry*, Cambridge Tracts in Mathematics, No. 29, 1933, or H. Whitney, *Differentiable manifolds*, Annals of Mathematics, vol. 37(1936), pp. 645–680.

given by $\theta^{-1}(\phi(t))$, if η is small enough. We say two parametrized curves starting at x_0 are *equivalent* if, when translated into E, they have the same tangent vector (in both magnitude and direction). Clearly the definition of equivalence is independent of the coördinate system chosen. Hence the classes of equivalent curves form a set of elements intrinsically defined in M; we call these *contravariant vectors* at x_0. Using a fixed coördinate system, we may obtain a (1-1) correspondence between contravariant vectors g at x_0 and vectors v in E at 0, merely by choosing, as an interval, the line segment of v, parametrized so that $t = 1$ at its end, and mapping it (or a portion of it, if it does not lie wholly in the region) into M with θ. We may *add* two contravariant vectors at x by taking the corresponding vectors in E, adding, and mapping back into M. Again the result is independent of the coördinate system chosen; hence *the contravariant vectors at x_0 form an intrinsically defined linear space*, the *tangent space $G(x_0)$ to M at x_0*.

We may obtain an intrinsic definition of $L(G(x_0)) = H(x_0)$ at x_0 by considering differentiable functions defined in a neighborhood of x_0, which vanish at x_0, and calling two functions equivalent if their partial derivatives at x_0 are the same in any coördinate system. To add covariant vectors, we need merely add the corresponding functions.

We shall consider briefly *covariant differentiation* in M. Suppose that to any two sufficiently near points x_0 and x_1 of M corresponds a linear map $\Psi_{x_1 x_0}$ of $G(x_1)$ into $G(x_0)$, so that certain simple continuity and linearity properties are satisfied, which we shall not make precise. This will define an *affine connection*[20] in M. Now let $A(x)$ be a differentiable tensor field, being, for each x, an element of $T(p, q; x)$ (using $G(x)$). Let g be any contravariant vector at x_0, and let C, given by $\phi(t)$, be a corresponding parametrized curve. Then if $x_t = \phi(t)$, we may define

$$\nabla_g A(x_0) = \lim_{t \to 0} \frac{1}{t} [\Psi_{x_t x_0} A(x_t) - A(x_0)]. \tag{12.1}$$

(Of course $\Psi_{x_t x_0}$ may be used to translate a tensor at x_t into a tensor at x_0.) For each g at x_0, $\nabla_g A(x_0)$ is a tensor of $T(p, q; x_0)$, and it depends linearly on g; hence we have a linear map of $G(x_0)$ into $T(p, q; x_0)$. By Theorem 25, there is a natural isomorphism

$$Ch_{T(p,q,x_0)}(G(x_0)) \approx T(p, q; x_0) \circ L(G(x_0)) \approx T(p, q + 1; x_0).$$

Hence, at each point x_0 we have a tensor of $T(p, q + 1; x_0)$, of the same contravariant order as A and of covariant order one greater; this is the *covariant derivative* of A at x_0. Again, the definition is intrinsic.

[20] By using a coördinate system about x_0 and letting $x_1 \to x_0$, we may use this connection to obtain an affine connection in the ordinary sense. Conversely, given an ordinary affine connection, we may define geodesics in M, and by following along them, define a connection as above. If we imbed M in a Euclidean space as in Whitney, loc. cit., Theorem 1, we may realize the tangent spaces by tangent planes of dimension n, and define an affine connection by projecting one tangent plane onto another.

13. **Products, general linear spaces.** A representation $\sum g_i \cdot h_i$ of an element α of $G \circ H$ is *minimal* if there is no representation with fewer summands. The *rank* $\rho(\alpha)$ of α is the number of summands in a minimal representation of α. We consider $0 = 0 \cdot 0$ as having no summands, and set $\rho(0) = 0$.

We collect some known results (at least for finite dimensional spaces) in the following theorem.

THEOREM 27. *$G \circ H$ is a linear space. For any α in $G \circ H$ there are corresponding linear subspaces $G(\alpha)$ and $H(\alpha)$ of G and H with the following properties.*

(a) *There is a representation $\sum g_i \cdot h_i$ for α with g_i in $G(\alpha)$, h_i in $H(\alpha)$. In any representation $\sum g_i' \cdot h_i'$ for α, $G(\alpha) \subset G(g_1', \cdots)$, $H(\alpha) \subset H(h_1', \cdots)$.*

(b) $\dim G(\alpha) = \dim H(\alpha) = \rho(\alpha)$; *$G(\alpha)$ and $H(\alpha)$ are $G(g_1, \cdots)$ and $H(h_1, \cdots)$ in any minimal representation $\sum g_i \cdot h_i$ of α.*

(c) *$\sum g_i \cdot h_i$ is minimal if and only if the sets $g_1, \cdots$ and $h_1, \cdots$ are each independent.*

(d) *If $g_1, \cdots, g_m$ and $h_1, \cdots, h_n$ are bases in subspaces G' of G and H' of H, and $\alpha = \sum a_{ij} g_i \cdot h_j$, then $\rho(\alpha) = \text{rank} \,\| a_{ij} \|$.*

(e) *If $g \cdot h = 0$, then either $g = 0$ or $h = 0$.*

(f) *$g \cdot h = g' \cdot h' \neq 0$ if and only if $g' = ag$, $h' = (1/a)g$ for some real a.*

The first statement follows from Theorems 1 and 13.

Suppose $\alpha = \sum g_i' \cdot h_i' = \sum g_i'' \cdot h_i''$, g_i' in G', g_i'' in G'', h_i' and h_i'' in H^*. Set $G^* = G' \cap G''$, and choose subspaces G_1 and G_2 (possibly containing 0 alone) such that

$$G' + G'' = G^* \oplus G_1 \oplus G_2, \qquad G_1 \subset G', \qquad G_2 \subset G''.$$

Choose bases $\{g_i^*\}$ in G^*, $\{g_i^1\}$ in G_1, $\{g_i^2\}$ in G_2; then all the g's form a base in $G' + G''$. By Theorem 23, we may write uniquely, for some h_i^*, etc., in H^*,

$$\alpha = \sum g_i^* \cdot h_i^* + \sum g_i^1 \cdot h_i^1 + \sum g_i^2 \cdot h_i^2.$$

Now $G' = G^* \oplus G_1$; hence, if we reduce $\sum g_i' \cdot h_i'$ to this normal form, the third group of terms will not appear. As the normal form is unique, the third sum $= 0$. Similarly, as $G'' = G^* \oplus G_2$, the second sum vanishes. Hence $\alpha = \sum g_i^* \cdot h_i^*$ can be expressed by using g's from $G' \cap G''$ alone. Hence there is a minimal subspace $G(\alpha)$ which may be used. Find similarly a minimal $H(\alpha)$. Now α can be expressed, by using $G(\alpha)$ and $H' \supset H(\alpha)$, and $G' \supset G(\alpha)$ and $H(\alpha)$. Choosing bases properly in G' and H' and using the first normal form, we see at once that α may be expressed, using $G(\alpha)$ and $H(\alpha)$. This proves (a).

Next we show that rank $\| a_{ij} \|$ depends on α alone. Suppose $\{g_i\}$ and $\{g_i'\}$ are bases in G', $\{h_i\}$ is a base in H', and $\alpha = \sum a_{ij} g_i \cdot h_j = \sum a_{kj}' g_k' \cdot h_j$. If $g_i = \sum_k b_{ki} g_k'$, then $a_{kj}' = \sum_i b_{ki} a_{ij}$, i.e., $A' = BA$. As B is non-singular, rank $A =$ rank A'. Similarly, a change of base in H' causes no change in the rank. If $G'' \supset G'$ and $H'' \supset H'$, and we choose bases in these spaces containing the above g_i and h_i, then $\sum a_{ij} g_i \cdot h_j$ is also a normal form for α, using G'' and H''. The new $\| a_{ij} \|$ is the old $\| a_{ij} \|$ with extra rows and columns of zeros; the ranks are therefore the same. Now given any two representations of α in normal

form, using the pair G', H' and the pair G'', H'', we may also write α in normal form, using $G' + G''$ and $H' + H''$. The above proof shows that all ranks of matrices are the same.

If $\sum_{i=1}^{\rho(\alpha)} g_i \cdot h_i$ is minimal, then obviously the sets $\{g_i\}$ and $\{h_i\}$ are independent. They form bases in spaces G' and H', say, and the expression $\sum g_i \cdot h_i$ is then in normal form. The matrix is the unit matrix, and hence is of rank $\rho(\alpha)$. This proves (d). As $G(\alpha) \subset G'$, and dim $G(\alpha) < \rho(\alpha)$ is clearly impossible, $G(\alpha) = G'$ and dim $G(\alpha) = \rho(\alpha)$; similarly for $H(\alpha)$. (b) is now proved. If $\alpha = \sum_{i=1}^{r} g_i \cdot h_i$ and the sets $\{g_i\}$, $\{h_i\}$ are independent, then we have a representation in normal form, with matrix of rank r; hence $r = \rho(\alpha)$, and $\sum g_i \cdot h_i$ is minimal. This proves (c).

To prove (e), suppose $g \neq 0$, $h \neq 0$. Then $g \cdot h$ is minimal, by (c), hence $\rho(g \cdot h) = 1$, and $g \cdot h \neq 0$. (f) follows from the fact that for $\alpha = g \cdot h = g' \cdot h' \neq 0$, $G(\alpha) =$ all multiples of $g =$ all multiples of g'.

THEOREM 28. *If G' is a linear subspace of G, then there is a natural isomorphism $G' \circ H \approx \sum^*(G' \cdot H)$.*

Using $\sum g_i \times h_i$ in $G' \circ H$, set $\phi(\sum g_i \times h_i) = \sum g_i \cdot h_i$. Clearly ϕ is a uniquely defined homomorphism onto the whole of $\sum^*(G' \cdot H)$. Suppose $\phi(\sum g_i \times h_i) = \sum g_i \cdot h_i$, $\sum g_i \times h_i \neq 0$. We may suppose $\sum g_i \times h_i$ is minimal. Then the sets $\{g_i\}$ and $\{h_i\}$ are independent, and hence $\sum g_i \cdot h_i$ is minimal, by the last theorem, and $\sum g_i \cdot h_i \neq 0$. Hence ϕ is (1-1), and this completes the proof.

14. **On topological linear spaces.** We shall use the following definition. If $G' \subset G^*$, a *projection* of G^* into G' is a linear map of G^* into G' such that every element of G' is fixed.

Definition. We shall call a *topological linear space* G a linear space with sets $U, V, \cdots$, called neighborhoods (of 0), such that:

(1) 0 is in every U;

(2) given U, V, there is a $W \subset U \cap V$;

(3) given U, there is a V such that for $-1 \leqq a \leqq 1$, $aV \subset U$;

(4) given U, there is a V with $V + V \subset U$;

(5) for every U and every g in G there is an a with g in aU;

(6) for every finite dimensional subspace G' of G and every natural neighborhood U' in the space G' (see §10), there is a neighborhood U in G with the following property. If $G^* \supset G'$ is a finite dimensional subspace, then there is a projection of G^* into G' which carries $U \cap G^*$ into U'.

We shall relate this definition to Definition 2b of von Neumann.[21]

Note that (6) implies a separation postulate: If $g \neq 0$, then there is a U which does not contain g. The reason for using our (6) is that with it one may prove the same property, and hence the separation postulate, in tensor products.

[21] J. von Neumann, *On complete topological spaces*, Transactions of the American Mathematical Society, vol. 37(1935), pp. 1–20. We refer to this paper as N.

Theorem 29. *A topological linear space (even with (6) replaced by a separation postulate) is a regular Hausdorff space; $g + g'$ and ag are each continuous in both variables.*

As our definition has all the properties in N, Definition 2b, except his (2) and (7), and a separation postulate holds, his proof holds without change.[22] We may now use U_{cl} = closure of U and U_i = inner points of U, etc., as in N.

Preparatory to proving Theorem 30, we note the following facts.

(a) If a set of sets U satisfies the above properties, then so do the sets U_{cl}, the sets U_i, and the sets $U - U$ (= all $g - g'$, g and g' in U).

(b) The sets U_{cl}, $(U_{cl})_i$, and $U - U$ define the same topology (i.e., give the same definition of S_i for any S) as the sets U.

These hold also if N, Definition 2b is used. To prove these facts, first note that N, Theorem 3, in particular, $(aS)_{cl} = aS_{cl}$, holds for closures; the proof is essentially the same. (a) and (b) now follow easily, using especially: $U_{cl} \subset U + U$; $V + V \subset U$ implies $V \subset U_i$.

(c) In a convex space as in N, we may suppose that the U are convex, and either closed or open, and that $-U = U$.

For we may use either the U_{cl} or the $(U_{cl})_i$. The U_{cl} are convex, by N, Theorem 13. To prove this for $(U_{cl})_i = S_i$, take g and g' in S_i and $0 < a < 1$. Set $g^* = ag + (1 - a)g'$, and choose V so that $g + V \subset S$, $g' + V \subset S$. Then

$$g^* + V \subset a(g + V) + (1 - a)(g' + V) \subset S_{cl} = S,$$

and hence g^* is in S_i. Finally, replace each U by $U - U = U'$; then all former properties hold, and $-U' = U'$.

Lemma 6. *Let G be a convex topological linear space as in N, satisfying our* (c). *Let $g_1, \cdots, g_\mu, g'$, be independent; let $g_1, \cdots, g_\mu$ determine the subspace G_1 of G, and the whole set, the subspace G^*. Let m be an integer $\leqq \mu$. Let U be a neighborhood such that*

$$\text{(14.1)} \qquad \textit{if } \sum_{i=1}^{\mu} a_i g_i \textit{ is in } U, \quad \textit{then} \quad |a_i| \leqq t_i \qquad (i = 1, \cdots, m).$$

Then there is a projection of G^* into G_1 such that the projection of $U \cap G^*$ satisfies the same inequalities.

It will not restrict the generality if we suppose that t_i are the smallest numbers such that (14.1) holds. As $-U = U$, no inequality in (14.1) can be bettered now.

First, suppose we have two elements

$$\text{(14.2)} \qquad g_1'' = \sum a_i g_i + cg', \qquad g_2'' = \sum b_i g_i + cg', \qquad \text{in } U;$$

then as $U = -U$ is convex,

$$\tfrac{1}{2}(g_1'' - g_2'') = \sum \tfrac{1}{2}(a_i - b_i)g_i \qquad \text{is in } U,$$

[22] In Hausdorff, *Mengenlehre*, Berlin, 1927, there is an error on p. 229: (5) does not follow from (6), as shown by a space in which the only open sets are the null set and the whole space. In N, proof of Theorem 6, one should mention that a separation postulate holds, as a consequence of Definition 2b, (2).

and hence

$$|a_i - b_i| \leqq 2t_i \qquad (i = 1, \cdots, m). \tag{14.3}$$

Now take any $c \geqq 0$ for which there is an element of the form (14.2) in U; let $\phi_i(c)$ and $\psi_i(c)$ $(i = 1, \cdots, m)$ be the greatest lower bound and least upper bound respectively of all numbers d such that for some numbers $a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_\mu$,

$$\sum_{j \neq i} a_j g_j + (\pm t_i + d)g_i + cg' \qquad \text{is in } U \quad (- \text{ for } \phi, + \text{ for } \psi).$$

In other words, $\phi_i(c)$ and $\psi_i(c)$ show how much U sticks out beyond the rectangle of the t_i, in the g_i direction, at the height c, with respect to the direction of g'. By the choice of the t_i, $\phi_i(0) = \psi_i(0) = 0$.

By (14.3), $\phi_i(c) \geqq \psi_i(c)$. We now show that

$$\text{if} \quad 0 < c < c', \quad \text{then} \quad \frac{\phi_i(c)}{c} \leqq \frac{\phi_i(c')}{c'}, \qquad \frac{\psi_i(c)}{c} \geqq \frac{\psi_i(c')}{c'}. \tag{14.4}$$

Suppose, for instance, the first inequality is false. Then there are numbers a_j $(j \neq i)$, e, such that

$$g_1' = \sum_{j \neq i} a_j g_j + (-t_i + d)g_i + c'g' \qquad \text{is in } U,$$

$$d = \frac{c'}{c}[\phi_i(c) - e], \qquad e > 0.$$

By the choice of the t_i, there is an element

$$g_2' = \sum_{j \neq i} b_j g_j + (-t_i + d')g_i \text{ in } U, \qquad d' < \frac{c'e}{c' - c}.$$

As U is convex,

$$\frac{c}{c'} g_1' + \frac{c' - c}{c'} g_2' = \sum_{j \neq i} a_j' g_j + \left(-t_i + \frac{c}{c'} d + \frac{c' - c}{c'} d'\right) g_i + cg'$$

is in U. But also

$$\frac{c}{c'} d + \frac{c' - c}{c'} d' = \phi_i(c) - e + \frac{c' - c}{c'} d' < \phi_i(c),$$

contradicting the definition of $\phi_i(c)$.

The inequalities show that we may define

$$\phi_i' = \lim_{c \to 0+} \frac{\phi_i(c)}{c}, \qquad \psi_i' = \lim_{c \to 0+} \frac{\psi_i(c)}{c}; \tag{14.5}$$

then

$$\frac{\phi_i(c)}{c} \geqq \phi_i' \geqq \psi_i' \geqq \frac{\psi_i(c)}{c} \qquad (c > 0; i = 1, \cdots, m). \tag{14.6}$$

Set

$$(14.7)\qquad g'' = g' + \sum_{i=1}^{m} \phi_i' g_i;$$

we shall show that if we project U along the direction of g'' into G_1, i.e., use $\phi(\sum a_i g_i + a''g'') = \sum a_i g_i$, then (14.1) will hold for the projection. As $-U = U$, it will be sufficient to consider the part of U with $c > 0$. If this is false, say for i, then there is an element

$$(14.8)\qquad \sum_{j=1}^{\mu} a_j g_j + cg'' \text{ in } U, \qquad c > 0, \qquad a_i > t_i \text{ or } a_i < -t_i.$$

Using (14.7), we have

$$(14.9)\qquad \sum_{j\neq i}^{j \leqq m} (a_j + c\phi_j')g_j + \sum_{j=m+1}^{\mu} a_j g_j + (a_i + c\phi_i')g_i + cg' \text{ in } U.$$

Suppose first that $a_i < -t_i$. Then

$$a_i + c\phi_i' < -t_i + c\,\frac{\phi_i(c)}{c} = -t_i + \phi_i(c),$$

contradicting the definition of $\phi_i(c)$. Next, if $a_i > t_i$, then

$$a_i + c\phi_i' > t_i + c\psi_i' \geqq t_i + \psi_i(c),$$

contradicting the definition of $\psi_i(c)$. This completes the proof.

THEOREM 30. *Any convex topological linear space as in* N *is a topological linear space as here defined, even if his* (2) *is replaced by a separation postulate.*

We may suppose his neighborhoods satisfy our (c). We must prove our (6). Let $g_1, \cdots, g_m$ form a base for G', and choose $t_1, \cdots, t_m$ so that all points $\sum a_i g_i$, $|a_i| \leqq t_i$, lie in U'. Let R be the closed region $|a_i| \leqq t_i$, and let A be its boundary. For each g in A, we may choose a $U(g)$ not containing it, and a $V(g)$ so that $V(g) - V(g) \subset U(g)$. As the operations in G are continuous, $\sum a_k g_k$ is continuous in the a_k, so the $(V(g))_i \cap G'$ (S_i = inner points of S) are open in the natural topology in G'; hence a finite number of the sets $g + V(g) \cap G'$ cover A. Let U be a neighborhood in the corresponding set $V(g_1) \cap \cdots \cap V(g_\lambda)$. Now U contains no element of A. For suppose g is in $A \cap U$. Say g is in $g_k + V(g_k)$. Then as g is in $U \subset V(g_k)$, g_k is in $V(g_k) - V(g_k) \subset U(g_k)$, a contradiction. As U is convex, $U \cap G'$ is in the complement of A in R.

Let $g_{m+1}', \cdots, g_n'$ form, with $g_1, \cdots, g_m$, a base in G^* (if $G^* \neq G'$), and let G_i be the space generated by $g_1, \cdots, g_i'$ $(i = m+1, \cdots, n)$. We shall prove, by induction on i, that G_i can be projected into G' so that $U \cap G_i$ goes into R; as $R \subset U'$, the case $i = n$ gives the theorem. There will be elements $g_{m+1}, \cdots, g_n$ such that $g_1, \cdots, g_i$ also determine G_i, and the projection of G_i into G' is with respect to $g_{m+1}, \cdots, g_i$:

$$\phi\left(\sum_{k=1}^{i} a_k g_k\right) = \sum_{k=1}^{m} a_k g_k.$$

Suppose we have found the elements $g_{m+1}, \cdots, g_\mu$. As the projection of G_μ into G' carries $U \cap G_\mu$ into R, (14.1) is satisfied. Hence we may apply Lemma 6 with $g' = g'_{\mu+1}$, $G^* = G_{\mu+1}$; this gives a projection of $G_{\mu+1}$ into G_μ such that the projection of $U \cap G_{\mu+1}$ satisfies (14.1). Let $g_{\mu+1}$ give the direction of the projection; then projecting $G_{\mu+1}$ into G' with respect to $g_{m+1}, \cdots, g_{\mu+1}$ carries $U \cap G_{\mu+1}$ into R, as required.

15. **Products, topological linear spaces.** We prove

THEOREM 31. *If G and H are topological linear spaces, so is $G \circ H$, the topology being given by* (10.12). *(We may use either open or closed neighborhoods in G and in H; see* §14.) *The multiplication $g \cdot h$ is continuous. The topology in $G \circ H$ depends only on the topologies in G and in H, not on the neighborhood systems employed.*

First, $G \circ H$ is a linear space, by Theorem 27. We shall prove the postulates of §14. (1) is trivial. To prove (2), take any two neighborhoods $N(U_1, \cdots; V_1, \cdots) = N(U_i; V_i)$, and $N(U'_i; V'_i)$. Choose U''_i in $U_i \cap U'_i$ and V''_i in $V_i \cap V'_i$; then

$$N(U''_i; V''_i) \subset N(U_i; V_i) \cap N(U'_i; V'_i).$$

To prove (3), given $N(U_i; V_i)$, choose U'_i so that $aU'_i \subset U_i$ if $|a| \leqq 1$ ($i = 1, 2, \cdots$). Then

$$aN(U'_i; V'_i) = N(aU'_i; V'_i) \subset N(U_i; V_i).$$

To prove (4), take any $N(U_i; V_i)$. Choose U'_i and V'_i so that

$$U'_i \subset U_{2i-1} \cap U_{2i}, \qquad V'_i \subset V_{2i-1} \cap V_{2i}.$$

Now take any $\sum g_i \cdot h_i$ and $\sum g'_i \cdot h'_i$ in $N(U'_i; V'_i)$. As $g_1 \cdot h_1$ is in $U'_1 \cdot V'_1 \subset U_1 \cdot V_1$, $g'_1 \cdot h'_1$ is in $U'_1 \cdot V'_1 \subset U_2 \cdot V_2$, $g_2 \cdot h_2$ is in $U'_2 \cdot V'_2 \subset U_3 \cdot V_3$, etc., we see that $\sum g_i \cdot h_i + \sum g'_i \cdot h'_i$ is in $N(U_i; V_i)$.

To prove (5), take any $\sum_{i=1}^{s} g_i \cdot h_i$ and any $N(U_i; V_i)$. Choose U^*_i and V^*_i so that

$$cU^*_i \subset U_i, \qquad cV^*_i \subset V_i, \qquad (|c| \leqq 1; i = 1, \cdots, s),$$

take a_i and b_i so that

$$g_i \text{ is in } a_i U^*_i, \qquad h_i \text{ is in } b_i V^*_i,$$

and let a be the largest of the $|a_i|$ and $|b_i|$. Then

$$a_i U^*_i = a(a_i/a) U^*_i \subset aU_i, \qquad \text{etc.};$$

it follows that $\sum g_i \cdot h_i$ is in

$$a^2 N(U_i; V_i) = N(aU_i; aV_i).$$

We now prove (6). Let F' be a subspace of $F = G \circ H$, generated by $f_1, \cdots, f_s$. Set (see Theorem 27)

$$G' = G(f_1) + \cdots + G(f_s), \qquad H' = H(f_1) + \cdots + H(f_s).$$

Let $g_1, \cdots, g_m$ and $h_1, \cdots, h_n$ be bases in G' and H'; set $f_{ij} = g_i \cdot h_j$. Then the f_{ij} form a base (see Theorem 26) in a space $F'' \supset F'$. Take a fixed projection of F'' into F'. Given a natural neighborhood N' in F', we may choose a natural neighborhood N'' in F'' which projects into a subset of N'. As any projection of an F^* into F'' combines with the above projection to give a projection of F^* into F', it is sufficient to prove (6) with F', N' replaced by F'', N''.

Choose $\epsilon > 0$ so that any $\sum a_{ij} f_{ij}$ with each $| a_{ij} | \leqq \epsilon$ is in N''. Let A and B be the sets of elements $\sum a_i g_i$ and $\sum b_i h_i$ in G' and H' with $| a_i | \leqq \frac{1}{2}\epsilon$, $| b_i | \leqq 1$. Choose U_1 in G by (6) so that any $U_1 \cap G^*$ can be projected into A, and choose V_1 in H so that any $V_1 \cap H^*$ can be projected into B. Choose $U_2, U_3, \cdots$ so $2U_i \subset U_{i-1}$, and set $V_2 = V_3 = \cdots = V_1$. Set $N = N(U_i; V_i)$.

Now take any $F^* \supset F''$. Choose a base $f_1^*, \cdots, f_t^*$ in F^*, and set $G^* = \sum G(f_i^*)$, $H^* = \sum H(f_i^*)$. Choose projections of G^* into G' and H^* into H' so $U_1 \cap G^*$ goes into A and $V_1 \cap H^*$ goes into B. If G_1 is the subset of G^* projecting into 0 in G', and $g_{m+1}, \cdots, g_\mu$ is a base in G_1, then $g_1, \cdots, g_\mu$ is a base in G^*; choose a base $h_1, \cdots, h_\nu$ in H^* similarly. Now any element of F^* can be written uniquely in the form

$$f = \sum_{(i,j)=(1,1)}^{(m,n)} a_{ij} f_{ij} + \sum{}' a_{ij} g_i \cdot h_j,$$

where in $\sum'$, either $i > m$ or $j > n$. (Not all such elements need be in F^*.) Dropping out the second group of terms defines a projection of F^* into F''.

We shall show by induction that any $(U_i \cdot V_i) \cap F^*$ projects into elements $\sum a_{kl} f_{kl}$ with $| a_{kl} | \leqq \epsilon/2^i$; it will follow that $N = \sum_i^* (U_i \cdot V_i)$ projects into N''.

Take first any α in $(U_1 \cdot V_1) \cap F^*$; we may suppose $\alpha \neq 0$. Then $\alpha = g \cdot h$, g in U_1, h in V_1. As α is in F^*, $G(\alpha) \subset G^*$. But also $G(\alpha) \subset G(g)$; hence $G(\alpha) \subset G^* \cap G(g)$. As $\alpha \neq 0$, $G(\alpha)$ contains elements $\neq 0$, which implies that g is in G^*. Similarly h is in H^*. Say

$$g = \sum_{i=1}^{\mu} a_i g_i, \qquad h = \sum_{j=1}^{\nu} b_j h_j.$$

Then as g projects into A and h into B, $g \cdot h$ projects into

$$\sum_{(i,j)=(1,1)}^{(m,n)} a_i b_j f_{ij}, \qquad | a_i b_j | < \tfrac{1}{2}\epsilon,$$

so that the statement holds for $(U_1 \cdot V_1) \cap F^*$. Supposing it holds for $k - 1$, we shall prove it for k. Take any g in U_k and h in V_k such that $g \cdot h$ is in F^*. Then

$$2g \text{ is in } 2U_k \subset U_{k-1}, \qquad h \text{ is in } V_{k-1},$$

so that $2(g \cdot h)$ is in $(U_{k-1} \cdot V_{k-1}) \cap F^*$, and hence projects into $\sum a_{ij} f_{ij}$ with $|a_{ij}| \leqq \epsilon/2^{k-1}$. Hence the required inequality on $g \cdot h$ holds. This completes the proof of (6).

To show that $g \cdot h$ is continuous, as $+$ is continuous in $G \circ H$ (Theorem 29), and

$$(15.1) \qquad (g + g') \cdot (h + h') - g \cdot h = g \cdot h' + g' \cdot h + g' \cdot h',$$

it is sufficient to show that $g \cdot h'$ is continuous in h' at $h' = 0$, $g' \cdot h$ is continuous in g' at $g' = 0$, and $g' \cdot h'$ is continuous in g' and h' at $g' = 0$, $h' = 0$. For the first case, given $N = N(U_i\,;\,V_i)$, choose a so that g is in aU_1, and choose V so that $aV \subset V_1$ (N, Theorem 1, with $n = 1$). Then

$$g \cdot V \subset aU_1 \cdot V = U_1 \cdot aV \subset U_1 \cdot V_1 \subset N.$$

The second case is similar. The third is clear, as $U_1 \cdot V_1 \subset N$.

Finally, let $\{U\}$, $\{\bar{U}\}$ and $\{V\}$, $\{\bar{V}\}$ be equivalent neighborhood systems in G and H, respectively. Given an $N(U_i\,;\,V_i)$, choose $\bar{U}_i \subset U_i$ and $\bar{V}_i \subset V_i$ $(i = 1, 2, \cdots)$; then $\bar{N}(\bar{U}_i\,;\,\bar{V}_i) \subset N(U_i\,;\,V_i)$. Similarly find an N in any $\bar{N}$. Hence the $\{N\}$ and $\{\bar{N}\}$ are equivalent. The theorem is now completely proved.

III. Topological groups

16. **The topological tensor product**. An Abelian *topological group* G is an Abelian group which is at the same time a Hausdorff space,[23] and in which $\phi(g, g') = g + g'$ and $\psi(g) = -g$ are continuous. If $U, U', \cdots$ are the neighborhoods of 0, we may let the sets $g + U, g + U', \cdots$ be the neighborhoods of the element g, without altering the topology. If we assume that the separation postulate in Hausdorff, page 229, (4), holds, then the postulate (5) follows.

We shall say the space G is *sequence-separable* if it contains a finite or denumerable set of points forming a dense set.[24]

If G and H are sequence-separable topological groups, we define their *topological tensor product* $G \circ H$, or *tensor product* simply, as follows. Let $\bar{g}_1, \bar{g}_2, \cdots$ and $\bar{h}_1, \bar{h}_2, \cdots$ be sequences of points dense in G and H, respectively. Let $P_1, P_2, \cdots$ be a sequence of pairs of elements, $P_i = (\bar{g}_{\mu_i}, \bar{h}_{\nu_i})$, such that each pair $(\bar{g}_j, \bar{h}_k)$ occurs infinitely often among the P_i. Let T' be the discrete tensor product of G and H, with elements $\sum g_i \times h_i$. For each sequence $U_1, U_2, \cdots$ of neighborhoods (of 0) in G and each sequence $V_1, V_2, \cdots$ in H, set

$$(16.1) \qquad \begin{aligned} Q'_i(U, V) &= \bar{g}_{\mu_i} \times V + U \times \bar{h}_{\nu_i} + U \times V, \\ N'(U_1, \cdots; V_1, \cdots) &= \sum_i{}^* Q'_i(U_i, V_i). \end{aligned}$$

Next, call two elements α, β of T' *equivalent*, $\alpha \sim \beta$, if every $\alpha + N$ contains β or vice versa, or if there is a succession $\alpha = \alpha_0, \alpha_1, \cdots, \alpha_n = \beta$, with α_i and

[23] See Hausdorff, loc. cit. Note that neighborhoods are open sets here.

[24] For metric spaces, this is the same as separability.

α_{i+1} equivalent as above. The sets of equivalent elements form the elements of the tensor product $T = G \circ H$. The *neighborhoods* N of 0 in $G \circ H$ are the images of the sets N' in T'; they are obtained by replacing $\times$ by $\cdot$ in (16.1). *Addition* in T is the image of addition in T'. The element $\sum g_i \cdot h_i$ in T is the image of $\sum g_i \times h_i$ in T'.

THEOREM 32. *$G \circ H$ is a sequence-separable Abelian topological group; the multiplication $g \cdot h$ satisfies* (1.1), *and is continuous. The topology in $G \circ H$ is independent of the sequences $\{\bar{g}_i\}$, $\{\bar{h}_i\}$, and of the neighborhood systems $\{U\}$, $\{V\}$, employed.*

We begin by showing that T' has all the properties of a topological group, except for the separation postulate. First we prove Hausdorff's postulates (B), (C) (loc. cit., p. 228). Given $N'(U_1, \cdots; V_1, \cdots) = N'(U_l; V_l)$ and $N'(U'_l; V'_l)$, take $U''_i \subset U_i \cap U'_i$ and $V''_i \subset V_i \cap V'_i$ $(i = 1, 2, \cdots)$; then clearly

$$N'(U''_l; V''_l) \subset N'(U_l; V_l) \cap N'(U'_l; V'_l). \tag{16.2}$$

To prove (C), it is sufficient to show that, for any $N'(U_l; V_l)$ and any α in $N'(U_l; V_l)$, there is an $N'(U'_l; V'_l)$ with

$$\alpha + N'(U'_l; V'_l) \subset N'(U_l; V_l). \tag{16.3}$$

As α is in $N'(U_l; V_l)$, it is in $\sum_{i=1}^{s} Q'_i(U_i; V_i)$ for some s. Choose numbers $\phi(1) > s, \phi(2) > \phi(1), \phi(3) > \phi(2), \cdots$ so that $P_i = P_{\phi(i)}$, and set $U'_i = U_{\phi(i)}$, $V'_i = V_{\phi(i)}$. Then

$$\begin{aligned} Q'_i(U'_i, V'_i) &= \bar{g}_{\mu_i} \times V'_i + U'_i \times \bar{h}_{\nu_i} + U'_i \times V'_i \\ &= \bar{g}_{\mu_{\phi(i)}} \times V_{\phi(i)} + U_{\phi(i)} \times \bar{h}_{\nu_{\phi(i)}} + U_{\phi(i)} \times V_{\phi(i)} = Q'_{\phi(i)}(U_{\phi(i)}, V_{\phi(i)}); \end{aligned}$$

hence $\sum_i^* Q'_i(U'_i, V'_i) \subset \sum_{i>s}^* Q'_i(U_i, V_i)$, and (16.3) follows.

We show that the group operations in T' are continuous. Given $N'(U_l; V_l)$, take

$$U'_i \subset U_i \cap (-U_i), \qquad V'_i \subset -V_i;$$

then

$$-Q'_i(U'_i, V'_i) = \bar{g}_{\mu_i} \times (-V'_i) + (-U'_i) \times \bar{h}_{\nu_i} + U'_i \times (-V'_i) \subset Q'_i(U_i, V_i);$$

hence

$$N'(U'_l; V'_l) \subset -N'(U_l; V_l), \tag{16.4}$$

and $-\alpha$ is continuous in α. To show that $\alpha + \beta$ is continuous, we must find $N'(U'_l; V'_l)$ corresponding to $N'(U_l; V_l)$ such that

$$N'(U'_l; V'_l) + N'(U'_l; V'_l) \subset N'(U_l; V_l). \tag{16.5}$$

Choose in succession integers

$$\phi(1), \qquad \psi(1) > \phi(1), \qquad \phi(2) > \psi(1), \qquad \psi(2) > \phi(2), \qquad \cdots$$

so that $P_i = P_{\phi(i)} = P_{\psi(i)}$. Take

$$U'_i \subset U_{\phi(i)} \cap U_{\psi(i)}, \qquad V'_i \subset V_{\phi(i)} \cap V_{\psi(i)}.$$

Then as $g_{\mu_{\phi(i)}} = g_{\mu_i}$, etc.,

$$Q'_i(U'_i, V'_i) \subset Q'_{\phi(i)}(U_{\phi(i)}, V_{\phi(i)}) \cap Q'_{\psi(i)}(U_{\psi(i)}, V_{\psi(i)}).$$

Hence

$$Q'_i(\cdots) + Q'_i(\cdots) \subset Q'_{\phi(i)}(\cdots) + Q'_{\psi(i)}(\cdots),$$

and (16.5) follows.

We now consider equivalent elements in T'. First we prove

(*) *If $\alpha \sim \beta$, then for every N', $\alpha + N'$ contains β.*

For suppose there is a succession $\alpha = \alpha_0, \alpha_1, \cdots, \alpha_n = \beta$, such that for each i, either every $\alpha_i + N'$ contains α_{i+1}, or every $\alpha_{i+1} + N'$ contains α_i. The latter condition implies the former. For given an N', choose $N'_1 \subset -N'$ by (16.4); then as α_i is in $\alpha_{i+1} + N'_1$, α_{i+1} is in $\alpha_i - N'_1 \subset \alpha_i + N'$. Next, given an N', choose N'_1 (using (16.5)) so that

$$N'_1 + N'_1 + \cdots + N'_1 \subset N' \qquad (n \text{ summands}).$$

Setting $A_k = N'_1 + \cdots + N'_1$ (k summands), we have

$$\alpha + N' \supset \alpha_0 + A_n \supset \alpha_1 + A_{n-1} \supset \cdots \supset \alpha_{n-1} + N'_1 \supset \beta,$$

as required.

Next we prove that T is a topological group. Let $\theta(\alpha)$ be the element of T containing the element α of T'. We must show that addition in T is uniquely defined; this is so if $\alpha \sim \alpha'$ and $\beta \sim \beta'$ imply $\alpha + \beta \sim \alpha' + \beta'$. Given any N', choose N'_1 so that $N'_1 + N'_1 \subset N'$. By the property (*), $\alpha + N'_1 \supset \alpha'$ and $\beta + N'_1 \supset \beta'$; hence

$$(\alpha + \beta) + N' \supset (\alpha + N'_1) + (\beta + N'_1) \supset \alpha' + \beta',$$

and $\alpha + \beta \sim \alpha' + \beta'$. Further,

$$\theta(\alpha + \beta) = \theta(\alpha) + \theta(\beta),$$

so that θ is a homomorphism of T' into T (which is clearly an Abelian group). To prove that T is a Hausdorff space, suppose $N_1 = \theta(N'_1)$ and $N_2 = \theta(N'_2)$ are given; take $N' \subset N'_1 \cap N'_2$; then $N = \theta(N') \subset N_1 \cap N_2$. Next, suppose α^* is in $N = \theta(N')$; then $\alpha^* = \theta(\alpha)$, α in N'. Choose N'_1 so that $\alpha + N'_1 \subset N'$; then $\alpha^* + \theta(N'_1) = \theta(\alpha + N'_1) \subset N$. To prove the separation postulate, suppose $\alpha^* \neq 0$. Say $\alpha^* = \theta(\alpha)$. As $\theta(0) = 0$, α is not ~ 0, and there is an N' not containing α. Set $N = \theta(N')$; then α^* is not in N. For if it were, then we would have $\alpha^* = \theta(\beta)$, β in N' and $\beta \sim \alpha$; but taking N'_1 by (16.3) so that $\beta + N'_1 \subset N'$, the property (*) gives $\alpha \subset \beta + N'_1 \subset N'$, a contradiction. To prove that the operations in T are continuous, given $N = \theta(N')$, take $N'_1 \subset -N'$; then $N_1 = \theta(N'_1) \subset -N$; also given $N = \theta(N')$, choose N'_1 so that $N'_1 + N'_1 \subset N'$; then $\theta(N'_1) + \theta(N'_1) \subset N$.

Next, (1) holds, as it holds for $\times$. We shall now show that $g \cdot h$ is continuous. First we show that it is continuous in h at $h = 0$; given $N = N(U_l; V_l)$, we must find V so that $g \cdot V \subset N$. As the $\bar{g}_i$ are dense in G, we may choose j so that $g - \bar{g}_{\mu_j}$ is in U_1. Choose $V \subset V_1 \cap V_j$; then

$$g \cdot V \subset (g - \bar{g}_{\mu_j}) \cdot V + \bar{g}_{\mu_j} \cdot V \subset U_1 \cdot V_1 + \bar{g}_{\mu_j} \cdot V_1 \subset N.$$

Similarly $g \cdot h$ is continuous in g at $g = 0$. Further, $g \cdot h$ is continuous in both variables at $g = 0$, $h = 0$; for $U_1 \cdot V_1 \subset N(U_l; V_l)$. Finally, as addition is continuous in T, (15.1) shows that $g \cdot h$ is continuous.

Next we show that T is sequence-separable; in fact, that the set of all finite sums $\sum \bar{g}_{p_i} \cdot \bar{h}_{q_i}$ is dense in T. As each element of T is a finite sum $\sum g_i \cdot h_i$ and $+$ is continuous, it is sufficient to show that for any $g \cdot h$ and any $N = N(U_l; V_l)$ there is a $\bar{g}_i \cdot \bar{h}_j$ in $g \cdot h + N$. As $\cdot$ is continuous, we may choose U and V so that $(g + U) \cdot (h + V) \subset g \cdot h + N$; we need now merely take $\bar{g}_i$ in $g + U$ and $\bar{h}_j$ in $h + V$.

That the topology in T is independent of the neighborhood systems chosen is trivial; see the end of the proof of Theorem 31. We must still show that the topology is independent of the choice of the $\bar{g}_i$ and $\bar{h}_i$. By symmetry, it is sufficient to show that if $\{\bar{g}_i\}$ is replaced by the dense sequence $\{g_i^*\}$, then any $N(U_l; V_l)$ contains an $N^*(U_l'; V_l')$, defined in terms of the g_i^*. Let ξ_i, η_i replace μ_i, ν_i. Given $N(U_l; V_l) = N_0$, choose $N_1, N_2, \cdots$ in succession so that $N_{i+1} + N_{i+1} \subset N_i$. As $\cdot$ and $+$ are continuous, we can choose U_i' and V_i' so that

$$Q_i^* = \bar{g}_{\xi_i} \cdot V_i' + U_i' \cdot \bar{h}_{\eta_i} + U_i' \cdot V_i' \subset N_i;$$

then for any s,

$$\sum_{i=1}^{s} Q_i^* \subset N_1 + \cdots + N_{s-1} + N_s \subset N_1 + \cdots + N_{s-2} + N_{s-1} + N_{s-1}$$

$$\subset N_1 + \cdots + N_{s-2} + N_{s-2} \subset \cdots \subset N_1 + N_1 \subset N_0,$$

and hence $N^*(U_l'; V_l') \subset N_0$, as required. This completes the proof of the theorem.

THEOREM 33. *Let $g_1^*, g_2^*, \cdots$ and $h_1^*, h_2^*, \cdots$ be (finite or infinite) sequences such that the sets $\sum a_i g_i^*$ and $\sum a_i h_i^*$ (integral a_i) are dense in G and H, respectively. Then we may use these sequences in place of dense sequences in defining the topology in $G \circ H$.*

Let $h_1, h_2, \cdots$ be either the above sequence $h_1^*, h_2^*, \cdots$, or a dense sequence in H. Arrange all $\sum a_i g_i^*$ in a sequence $\bar{g}_1, \bar{g}_2, \cdots$. Let $N(U_l; V_l)$ be defined in terms of the sets $\{\bar{g}_i\}$, $\{h_i\}$, and $N^*(U_l; V_l)$, in terms of the sets $\{g_i^*\}$, $\{h_i\}$. It is sufficient to show that these two sets of neighborhoods give the same topology in T. As the g_i^* occur among the $\bar{g}_i$, it is clear that any N contains an N^*; we must prove the converse.

Let $P_i = (\bar{g}_{\mu_i}, h_{\nu_i})$ and $P_i^* = (g_{\xi_i}^*, h_{\eta_i})$ define the sequences of pairs defining

the N and the N^*. If $\bar{g}_{\mu_i} = \sum_j a_{ij} g_j^*$, set $m(i) = \sum_j |a_{ij}|$. Then $\bar{g}_{\mu_i} \cdot V$ is contained in $m(i)$ terms of the form $g_j^* \cdot (\pm V)$, and $Q_i(U'_i, V'_i)$ is contained in $m(i) + 2$ terms of forms appearing (except for the $\pm$) in $N^*(U_l; V_l)$. For each i we may choose $m(i) + 2$ numbers $\phi_1(i), \cdots, \phi_{m(i)+2}(i)$ such that $\phi_k(i) \neq \phi_l(j)$ whenever $i \neq j$, and the k-th part into which Q_i is split corresponds to part of $Q^*_{\phi_k(i)}$. Then if the U'_i and V'_i are chosen small enough, we will have

$$Q_i(U'_i, V'_i) \subset \sum_{k=1}^{m(i)} Q^*_{\phi_k(i)}(U_{\phi_k(i)}, V_{\phi_k(i)}),$$

and hence $N(U'_l; V'_l) \subset N^*(U_l; V_l)$, as required.

17. **Relation to linear spaces; examples.** We shall show that whenever the definitions of tensor products in Parts II and III both apply, they coincide.

THEOREM 34. *If G and H are sequence-separable topological linear spaces, then their topological tensor product T is the same as their topological reduced tensor product T^*.*

Let $\sum g_i \times h_i$ and $\sum g_i \cdot h_i$ denote elements of T^* and T, respectively. Set $\phi(\sum g_i \times h_i) = \sum g_i \cdot h_i$; we shall show that ϕ is a topological isomorphism. To show that ϕ is uniquely defined, we must show that $ag \cdot h = g \cdot ah$ for any real a; but this follows from the continuity of $g \cdot h$ (see §10). ϕ is a homomorphism; we shall show that it is continuous. Use $N(U_l; V_l)$ in T and $N^*(U_l; V_l)$ in T^*. Given $N(U_l; V_l)$, we wish to find $N^*(U'_l; V'_l)$ mapping into it. From (10.12) and (16.1) it is apparent that we may use $U'_i = U_i$, $V'_i = V_i$.

Next we show that for any $N^* = N^*(U_l; V_l)$, there is an $N = N(U'_l; V'_l) \subset \phi(N^*)$. Say $\bar{g}_1, \bar{g}_2, \cdots$ and $\bar{h}_1, \bar{h}_2, \cdots$ are the dense sequences used in G and H, and $P_1, P_2, \cdots, P_i = (\bar{g}_{\mu_i}, \bar{h}_{\nu_i})$, the pairs. By §14, (5), we may choose for each i a number a_i such that $\bar{g}_{\mu_i}$ is in $a_i U_{3i-2}$, and a number b_i such that $\bar{h}_{\nu_i}$ is in $b_i V_{3i-1}$. By von Neumann, loc. cit., Theorem 1 (with $n = 1$), there is a V''_i such that $a_i V''_i \subset V_{3i-2}$, and a U''_i such that $b_i U''_i \subset U_{3i-1}$. Choose

$$U'_i \subset U''_i \cap U_{3i}, \qquad V'_i \subset V''_i \cap V_{3i}.$$

Now

$$\begin{aligned}\bar{g}_{\mu_i} \cdot V'_i + U'_i \cdot \bar{h}_{\nu_i} + U'_i \cdot V'_i &\subset a_i U_{3i-2} \cdot \frac{1}{a_i} V_{3i-2} + \frac{1}{b_i} U_{3i-1} \cdot b_i V_{3i-1} + U_{3i} \cdot V_{3i} \\ &= \phi(U_{3i-2} \times V_{3i-2} + U_{3i-1} \times V_{3i-1} + U_{3i} \times V_{3i});\end{aligned}$$

hence $N \subset \phi(N^*)$.

Clearly ϕ maps T^* into the whole of T. When we have shown that ϕ is (1-1), the proof will be completed. Let T' be the discrete tensor product of G and H; use $g \circ h$ here. Take any $\alpha^* = \sum g_i \times h_i \neq 0$ in T^*; then $\alpha' = \sum g_i \circ h_i$ is a corresponding element of T'. There is an $N^* = N^*(U_l; V_l)$ which does not contain α^*. Construct $U'_1, U'_2, \cdots$ and $V'_1, V'_2, \cdots$ by the method given above, and set

$$N' = \sum_i{}^* (\bar{g}_{\mu_i} \circ V'_i + U'_i \circ \bar{h}_{\nu_i} + U'_i \circ V'_i).$$

The map $\psi(\sum g_i' \circ h_i') = \sum g_i' \times h_i'$ of T' into T^* is uniquely defined. By the choice of the U_i' and V_i', $\psi(N') \subset N^*$; hence α' is not in N'. Therefore, by the property (*) in §16, α' is not ~ 0, so that the corresponding element $\sum g_i \cdot h_i$ of T is $\neq 0$. Consequently $\alpha^* \neq 0$ implies $\phi(\alpha^*) \neq 0$, and ϕ is (1-1).

Examples. That the topology in (10.12) cannot be used in the general case is shown by the example $I_0 \circ Rl$. A neighborhood U of 0 in I_0 is 0 itself; hence $U \cdot V = 0 \cdot V = 0$, and $I_0 \circ Rl$ would be discrete; the multiplication $a \cdot g$ would not be continuous. However, the sets $1 \cdot V$ form a neighborhood system in $I_0 \circ Rl$. In fact, *if G has a finite number of generators $\bar{g}_1, \cdots, \bar{g}_n$, then the sets*

$$\bar{g}_1 \cdot V_1 + \cdots + \bar{g}_n \cdot V_n \qquad (V_1, \cdots, V_n \text{ neighborhoods in } H) \tag{17.1}$$

form a neighborhood system in $G \circ H$. This is an easy consequence of Theorem 33 (compare Theorem 26).

HARVARD UNIVERSITY.

THEOREMS ON RIESZ MEANS

By H. L. Garabedian and W. C. Randels

1. **Introduction.** We are concerned in this paper with means of the type

$$\sigma_n = \frac{p_0 s_0 + p_1 s_1 + \cdots + p_n s_n}{P_n}, \tag{1.1}$$

where

$$P_n = p_0 + p_1 + \cdots + p_n ,$$

$\{s_n\}$ is a given sequence, and $\sum p_n$ is a divergent series of positive terms. Since means of the type (1.1) were used in the early development of the Riesz typical means, they are called Riesz means and are designated by (R, p_n).[1] It is of particular interest to notice that all of the Riesz means of the type under consideration constitute regular definitions of summation.

2. **Relative inclusion.** Let us first consider a theorem relating to the relative inclusiveness of these methods.

Theorem 1. *A necessary and sufficient condition that* $(R, q_n) \subset (R, p_n)$ *is*[2] *that*

$$\frac{1}{P_n} \sum_{\nu=0}^{n-1} \left| \frac{p_\nu}{q_\nu} - \frac{p_{\nu+1}}{q_{\nu+1}} \right| Q_\nu + \frac{p_n Q_n}{q_n P_n} < N, \tag{2.1}$$

where N is independent of n.

This is an improvement on a theorem due to Cesàro[3] and Hardy.[4] Cesàro proved that $(R, q_n) \subset (R, p_n)$ provided that

$$\frac{p_n}{q_n} \geqq \frac{p_{n+1}}{q_{n+1}}, \tag{2.2}$$

while Hardy proved that the same result obtains when

$$\begin{cases} \text{(a)} \quad \dfrac{p_n}{q_n} \leqq \dfrac{p_{n+1}}{q_{n+1}}, \\ \text{(b)} \quad \dfrac{p_n Q_n}{q_n P_n} = O(1). \end{cases} \tag{2.3}$$

Received March 10, 1938.

[1] G. H. Hardy, Proc. London Math. Soc., (2), vol. 8(1910), pp. 301–320.

[2] We understand $A \subset B$ to mean that every sequence summable by a method A is also summable by a method B to the same limit, or that the method B includes the method A. Further, we interpret $A = B$ to mean that the two methods are equivalent, that is, each method includes the other.

[3] Cesàro, Bulletin des Sciences Mathématiques, (2), vol. 13(1889), pp. 51–54.

[4] Hardy, Quarterly Journal, vol. 38(1907), pp. 269–288.

We observe that (2.3a) implies

$$\frac{q_n}{p_n} \geqq \frac{q_{n+1}}{p_{n+1}},$$

and hence, by (2.2), that $(R, p_n) \subset (R, q_n)$. Accordingly, (2.3) is a sufficient condition for the equivalence of two methods of summation of the type being considered. Some examples of the inclusion relations implied by this argument and Theorem 1 are[5]

$$\begin{aligned}(2.4)\qquad &(R, 1/(n \log n \log_2 n)) \supset (R, 1/(n \log n)) \supset (R, 1/n) \supset (R, 1)\\ &= (C, 1) \approx (R, n^{\alpha}) \approx (R, n \log n) \supset (R, k^n) \quad (\alpha > -1), (k > 1).\end{aligned}$$

The case (R, n^{-1}), where one writes $P_n = \log n$, is called *logarithmic summability*, and has been the source of a number of investigations relating to the summability and convergence of slowly oscillating series.[6]

We start the proof of Theorem 1 by setting

$$(2.5)\qquad \rho_n = \frac{q_0 s_0 + q_1 s_1 + \cdots + q_n s_n}{Q_n}.$$

From (2.5) we have

$$s_\nu = \frac{\rho_\nu Q_\nu - \rho_{\nu-1} Q_{\nu-1}}{q_\nu}.$$

Then

$$\sigma_n = \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu s_\nu = \frac{1}{P_n} \sum_{\nu=0}^{n} \frac{p_\nu}{q_\nu} (\rho_\nu Q_\nu - \rho_{\nu-1} Q_{\nu-1}),$$

or

$$(2.6)\qquad \sigma_n = \frac{1}{P_n} \sum_{\nu=0}^{n-1} \left[\frac{p_\nu}{q_\nu} - \frac{p_{\nu+1}}{q_{\nu+1}}\right] \rho_\nu Q_\nu + \frac{p_n Q_n \rho_n}{q_n P_n}.$$

Now, $\lim \rho_n = s$ implies that $\lim \sigma_n = s$ if and only if the transformation from the sequence $\{\rho_n\}$ to the sequence $\{\sigma_n\}$ defined by (2.6) is regular. Applying the Silverman-Toeplitz conditions to this transformation, we find that it is regular if and only if

$$\frac{1}{P_n} \sum_{\nu=0}^{n-1} \left|\frac{p_\nu}{q_\nu} - \frac{p_{\nu+1}}{q_{\nu+1}}\right| Q_\nu + \frac{p_n Q_n}{q_n P_n} < N,$$

where N is independent of n. The remaining conditions for regularity may be checked without difficulty.

3. **Range of effectiveness**. The theorems which follow throw further light on the properties of the methods of Riesz means.

[5] Hardy, London Mathematical Society, (2), vol. 15(1916–17), pp. 72–80.
[6] See, for example, Hardy, loc. cit., footnote 1.

THEOREM 2. *If the sequence $\{s_n\}$ is summable (R, p_n) to the value s, then*[7]

$$s_n = o(P_n/p_n) + s.$$

As a result of this theorem we notice that if $P_n/p_n = O(1)$, then summability (R, p_n) is equivalent to convergence. For example, summability (R, k^n), for $k > 1$, is equivalent to convergence.[8]

We notice also that if a sequence $\{s_n\}$ is summable by methods for which $p_n = 1/(n \log n)$, $1/n$, $1/\log n$, then we have respectively $s_n = o(n \log n \log_2 n)$, $o(n \log n)$, $o(n)$. Moreover, it is known that for the Cesàro methods, (C, α), $s_n = o(n^\alpha)$, and that these estimates are the best possible. Therefore, none of the methods considered can include summability (C, α) for $\alpha > 1$. On the other hand, it follows from an example given by Hardy[9] that summability (C, α), for α arbitrarily large, cannot include summability (R, n^{-1}).

THEOREM 3. *A necessary and sufficient condition that a series $\sum u_n$, with partial sums s_n, which is summable (R, p_n) to the value s, should converge to s is that*

$$\sum_{\nu=1}^{n} P_{\nu-1} u_\nu = o(P_n). \tag{3.1}$$

If $s_n \to s$, then $\sigma_n - s_n \to 0$ as $n \to \infty$. Thus,

$$\sigma_n - s_n = \frac{p_0 s_0 + p_1 s_1 + \cdots + p_n s_n}{P_n} - s_n = o(1),$$

or

$$\sum_{\nu=1}^{n} p_{\nu-1}(s_n - s_{\nu-1}) = o(P_n). \tag{3.2}$$

Reordering the terms in (3.2) we have

$$\sum_{\nu=1}^{n} P_{\nu-1} u_\nu = o(P_n).$$

Since the steps in this proof are reversible, the sufficiency part of the proof is also established.

4. **Certain limitations**. In connection with Theorem 3 we observe that, if summability (R, p_n) implies that

$$\sum_{\nu=1}^{n} p_\nu u_\nu = o(P_n), \tag{4.1}$$

the addition of (3.1) and (4.1) gives the relation

$$\sum_{\nu=1}^{n} P_\nu u_\nu = o(P_n), \tag{4.2}$$

[7] Hardy and Riesz, *The General Theory of Dirichlet's Series*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 18, Theorem 21.

[8] Loc. cit., footnote 7, p. 35.

[9] Loc. cit., footnote 1.

which might be used in Theorem 3. This question is answered in a corollary to the following theorem.

THEOREM 4. *A necessary and sufficient condition that a single term may be removed from the beginning of an (R, p_n) summable series and that the series will remain (R, p_n) summable is that*

$$\frac{1}{P_n}\sum_{\nu=1}^{n}\left|\frac{p_{\nu-1}}{p_\nu}-\frac{p_\nu}{p_{\nu+1}}\right|P_\nu+\frac{p_nP_{n+1}}{p_{n+1}P_n}<N, \tag{4.3}$$

where N is independent of n. The series will be related as if convergent.

We are concerned with the series

$$u_0+u_1+u_2+\cdots, \tag{4.4}$$

$$u_1+u_2+u_3+\cdots, \tag{4.5}$$

which have respectively the partial sums

$$s_n=u_0+u_1+\cdots+u_n,$$

$$s_n'=u_1+u_2+\cdots+u_{n+1}.$$

Let the Riesz means for the series (4.4) and (4.5) be respectively σ_n and σ_n'.

We have

$$\sigma_n'=\frac{1}{P_n}\sum_{\nu=0}^{n}p_\nu s_\nu'=\frac{1}{P_n}\sum_{\nu=0}^{n}p_\nu s_{\nu+1}-u_0.$$

Using the technique employed in proving Theorem 1, we have

$$u_0+\sigma_n'=\frac{1}{P_n}\sum_{\nu=0}^{n}p_\nu\frac{P_{\nu+1}\sigma_{\nu+1}-P_\nu\sigma_\nu}{p_{\nu+1}},$$

or

$$u_0+\sigma_n'=\frac{1}{P_n}\sum_{\nu=1}^{n}\left[\frac{p_{\nu-1}}{p_\nu}-\frac{p_\nu}{p_{\nu+1}}\right]P_\nu\sigma_\nu+\frac{p_nP_{n+1}\sigma_{n+1}}{p_{n+1}P_n}-\frac{p_0P_0\sigma_0}{P_np_1}. \tag{4.6}$$

Now, $\lim \sigma_n=s$ implies that $\lim \sigma_n'=s-u_0$ if and only if the transformation from the sequence $\{\sigma_n\}$ to the sequence $\{\sigma_n'\}$ defined by (4.6) is regular. Applying the Silverman-Toeplitz conditions to this transformation, we find that it is regular if and only if

$$\frac{1}{P_n}\sum_{\nu=1}^{n}\left|\frac{p_{\nu-1}}{p_\nu}-\frac{p_\nu}{p_{\nu+1}}\right|P_\nu+\frac{p_nP_{n+1}}{p_{n+1}P_n}<N,$$

where N is independent of n. The remaining conditions for regularity may be easily checked.

We notice that the condition of Theorem 4 is satisfied if p_n/p_{n+1} steadily decreases or steadily increases. Thus, all of the special methods thus far considered have the property embodied in the statement of Theorem 4.

We turn now to a consideration of a corollary to Theorem 4.

COROLLARY 1. *If a series is* (R, p_n) *summable, a necessary and sufficient condition that*

$$\sum_{\nu=0}^{n} p_\nu u_{\nu+1} = o(P_n)$$

is that condition (4.3) *of Theorem* 4 *obtains.*

We have

$$\sigma_n' = \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu s_\nu' = \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu s_{\nu+1} - u_0$$

$$= \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu s_\nu + \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu u_{\nu+1} - u_0,$$

or

$$\sigma_n' = \sigma_n - u_0 + \frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu u_{\nu+1}.$$

The proof of the corollary follows immediately from this statement.

In connection with the conjecture raised after the proof of Theorem 3 let us prove that the condition

$$\sum_{\nu=0}^{n} p_\nu u_{\nu+1} = o(P_n) \tag{4.7}$$

implies that

$$\sum_{\nu=1}^{n} p_\nu u_\nu = o(P_n). \tag{4.8}$$

If we require that $\lim p_{n+1}/p_n = 1$, it certainly follows from (4.7) that $u_n = o(P_n/p_n)$ and that the difference

$$\frac{1}{P_n} \sum_{\nu=0}^{n} p_\nu u_{\nu+1} - \frac{1}{P_n} \sum_{\nu=1}^{n} p_\nu u_\nu = \frac{1}{P_n} \sum_{\nu=0}^{n-1} (p_\nu - p_{\nu+1}) u_{\nu+1} + \frac{p_n u_{n+1}}{P_n} = o(1). \tag{4.9}$$

Moreover, (4.7) and (4.9) imply (4.8).

We conclude that if $\lim p_{n+1}/p_n = 1$, a property fulfilled by all of the special methods of summation listed in (2.4), then condition (3.1) of Theorem 3 may be replaced by the condition

$$\sum_{\nu=1}^{n} P_\nu u\nu = o(P_n).$$

NORTHWESTERN UNIVERSITY.

SIMPLE LIE ALGEBRAS OVER A FIELD OF CHARACTERISTIC ZERO

By N. Jacobson

The present paper gives a resumé and extension of the theory of simple Lie algebras over an arbitrary field Φ of characteristic 0 developed in recent papers by Landherr and by the author.[1] The extension consists in part in dropping the restriction of normality. Isomorphisms and automorphisms of simple but not necessarily normal simple Lie algebras are considered. We have also discussed the problem of cogredience of matrices arising in this connection and in the last sections have considered in detail the theory for a real closed field and sketched the theory also for $\mathfrak{p}$-adic fields. In the latter discussion we have confined ourselves to referring to results in the literature that are applicable here and have merely supplemented these with the theory of Hermitian and skew-Hermitian matrices having elements in a quaternion algebra.

1. **Preliminaries.** If $\mathfrak{A}$ is an associative algebra over a field Φ, it is readily seen that the elements of $\mathfrak{A}$ constitute a Lie algebra $\mathfrak{A}_l$ relative to the composition $[a, b] = ab - ba$. Evidently if $\mathfrak{A} \cong \mathfrak{B}$, $\mathfrak{A}_l \cong \mathfrak{B}_l$ and if $a \rightarrow b$ is an anti-isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$, then $a \rightarrow -b$ is an isomorphism between $\mathfrak{A}_l$ and $\mathfrak{B}_l$. In particular if S is an automorphism (anti-automorphism) in $\mathfrak{A}$, then S ($-S$: $a \rightarrow -a^S$) is an automorphism in $\mathfrak{A}_l$. It is clear that the elements left invariant by an automorphism in a Lie algebra form a subalgebra. Hence if S is an anti-automorphism in $\mathfrak{A}$, the set $\mathfrak{S}_S$ of S-skew elements b ($b^S = -b$) form a Lie subalgebra of $\mathfrak{A}_l$.

The anti-automorphisms S and T of $\mathfrak{A}$ over Φ are *cogredient* if there exists an automorphism G of $\mathfrak{A}$ over Φ, such that $T = G^{-1}SG$. In this case G maps $\mathfrak{S}_S$ on $\mathfrak{S}_T$ and hence G is an isomorphism between these Lie algebras. If $S = T$, i.e., $SG = GS$, then G induces an automorphism in $\mathfrak{S}_S$. If G is inner, say $a^G = g^{-1}ag$, the condition $GS = SG$ is equivalent to $gg^S \in$ the centrum of $\mathfrak{A}$.

If $\mathfrak{L}$ is any (Lie) subalgebra of $\mathfrak{A}_l$, we define the *enveloping ring* $\mathfrak{R}$ of $\mathfrak{L}$ in $\mathfrak{A}$ to be the smallest subring of $\mathfrak{A}$ containing all the elements of $\mathfrak{L}$. $\mathfrak{R}$ is clearly the totality of elements of the form $\sum l_1 l_2 \cdots l_k$, where $l_i \in \mathfrak{L}$. Since $l\alpha \in \mathfrak{L}$ for any l in $\mathfrak{L}$ and any α in Φ, it is evident that $\mathfrak{R}$ is an algebra over Φ.

If $\mathfrak{L}$ is an arbitrary Lie algebra over Φ, it is well known that the correspondence between the element a in $\mathfrak{L}$ and the linear transformation $\mathbf{A}$ defined by $x\mathbf{A} \equiv [x, a]$ for x variable in $\mathfrak{L}$ is a representation, called the *adjoint representation*, of $\mathfrak{L}$:

Received March 11, 1938; presented to the American Mathematical Society, April 15, 1938.

[1] Landherr [1], Jacobson [1] and [2]. Numbers in brackets refer to the bibliography at the end of the paper.

if $a \rightarrow \mathbf{A}$, $b \rightarrow \mathbf{B}$, then $a + b \rightarrow \mathbf{A} + \mathbf{B}$, $a\alpha \rightarrow \mathbf{A}\alpha$ and $[a, b] \rightarrow [\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. Let $\mathfrak{R}$ be the enveloping ring of the linear transformation $\mathbf{A}$ in the ring of all linear transformations. If $\mathfrak{L}$ is simple, it can be shown that the centrum of $\mathfrak{R}$ is an algebraic field Σ containing Φ and $\mathfrak{R} \cong \Sigma_m$ the m-rowed matrix ring[2] with elements in Σ.[3] Σ is called the *extended centrum* of $\mathfrak{L}$. $\mathfrak{L}$ may be regarded as an algebra over Σ and when this is done, it becomes *normal simple*, i.e., the extension ($\mathfrak{L}$ over $\Sigma)_\Omega$ is simple for Ω the algebraic closure of Σ. Moreover, Σ is the only field of operators $\supset \Phi$ relative to which $\mathfrak{L}$ is closed and normal simple.

If G is an isomorphism between the simple Lie algebras $\mathfrak{L}_1$ and $\mathfrak{L}_2$ over Φ, it induces an isomorphism between their extended centrums Σ_1 and Σ_2. Hence if we identify Σ_1 and Σ_2 as Σ by means of some fixed isomorphism (not necessarily G), $\mathfrak{L}_1$ and $\mathfrak{L}_2$ may be regarded as normal simple algebras over Σ. Then G induces an automorphism in Σ such that $(a_1\xi)^G = a_1^G\xi^G$ for a_1 in $\mathfrak{L}_1$ and ξ in Σ. If Ω is any over-field of Σ, ($\mathfrak{L}_1$ over $\Sigma)_\Omega \cong$ ($\mathfrak{L}_2$ over $\Sigma)_\Omega$. By identifying $\mathfrak{L}_1$ and $\mathfrak{L}_2$ as $\mathfrak{L}$ we note that an automorphism G of $\mathfrak{L}$ over Φ defines an automorphism in Σ such that $(a\xi)^G = a^G\xi^G$. Hence if $\mathfrak{G}$ is the group of automorphisms of $\mathfrak{L}$ over Σ, $\mathfrak{G}_0$ the group of automorphisms of $\mathfrak{L}$ over Σ and $\mathfrak{T}$ the Galois group of Σ over Φ, then $\mathfrak{G}_0$ is invariant in $\mathfrak{G}$ and $\mathfrak{G}/\mathfrak{G}_0 \cong \mathfrak{T}_0$ a subgroup of $\mathfrak{T}$.

2. **Simple Lie algebras over an algebraically closed field.** From now on we suppose that Φ has characteristic 0 and in this section that Φ is also algebraically closed.

It is well known that the associative algebra Φ_n of n-rowed matrices ($n > 1$) with elements in Φ is simple and the derived algebra[4] Φ'_{nl} of Φ_{nl} is a simple Lie algebra of order $n^2 - 1$ over Φ. Φ'_{nl} may be defined also as the totality of matrices of trace 0 in Φ_{nl}. $1 \,\epsilon'\, \Phi'_{nl}$, where ϵ' means "is not an element of", and hence if $A_1, \cdots, A_m$, $m = n^2 - 1$, is a basis for Φ'_{nl}, $A_0, \cdots, A_m$ where $A_0 = 1$ is a basis for Φ_{nl} (or Φ_n). The enveloping ring of Φ'_{nl} in Φ_n is Φ_n. The algebras Φ'_{nl} for $n = 2, 3, \cdots$ constitute Cartan's class A. It has been shown by A. Weinstein[5] that if G is an automorphism of Φ'_{nl} over Φ, then there is a non-singular matrix G such that either $A^G = G^{-1}AG$ or $A^G = -G^{-1}A'G$ for all A in Φ'_{nl} and A' the transpose of the matrix A.

If n is odd it can be shown[6] that any involutorial anti-automorphism (i.a.a.) in Φ_n over Φ is cogredient to the i.a.a. $A \rightarrow A'$. The skew elements relative to this i.a.a. are the ordinary skew symmetric matrices, and if $n = 3, 5, 7, \cdots$ these form a simple Lie algebra $\mathfrak{S}$ of order $\frac{1}{2}n(n-1)$ (Cartan's class B). The

[2] In general, if $\mathfrak{A}$ is a ring (algebra), we denote the r-rowed matrix ring (algebra) with elements in $\mathfrak{A}$ by $\mathfrak{A}_r$.

[3] See Jacobson [3] for the results of this paragraph and the next.

[4] The derived algebra $\mathfrak{L}' = [\mathfrak{L}, \mathfrak{L}]$ of a Lie algebra $\mathfrak{L}$ is the totality of sums of commutators $[x, y]$, x, y in $\mathfrak{L}$. The algebras Φ'_{nl} belong to one of Cartan's classes of simple algebras (cf. Cartan [1], Chapter 5).

[5] Weinstein [1].

[6] The results quoted in this paragraph and the next may be found in Jacobson [1].

enveloping ring of $\mathfrak{S}$ in Φ_n is Φ_n. If G is an automorphism of $\mathfrak{S}$ over Φ and $n > 5$, there exists a matrix G such that $GG' = \gamma \neq 0$ in Φ and $A^G = G^{-1}AG$ for all A in $\mathfrak{S}$.

If n is even ($= 2\nu$), any i.a.a. in Φ_n is cogredient either to $A \to A'$ or to $A \to Q^{-1}A'Q$, where

$$Q = \begin{pmatrix} 0 & 1_\nu \\ -1_\nu & 0 \end{pmatrix}. \tag{1}$$

In the first case the skew elements are ordinary skew matrices and for $n = 6, 8, \cdots$ these form a simple Lie algebra of order $\frac{1}{2}n(n-1)$ (Cartan's class D). The automorphisms of $\mathfrak{S}$ if $n \neq 8$ have the same form as for the algebras of class B. In the second case for $n = 2, 4, \cdots$ we obtain a simple Lie algebra $\mathfrak{S}$ of order $\frac{1}{2}n(n+1)$ consisting of the matrices A such that $Q^{-1}A'Q = -A$ (Cartan's class C). The automorphisms of $\mathfrak{S}$ over Φ have the form $A^G = G^{-1}AG$, where G is a matrix such that $Q^{-1}G'QG = \gamma \neq 0$ or $G'QG = Q\gamma$. In either of the two cases the enveloping ring of $\mathfrak{S}$ is Φ_n.

Suppose now that $n > 5$ for class B, > 2 for class C and > 6 for class D. By a fundamental result of Cartan's[7] the algebras enumerated in this section and subject to these restrictions are not isomorphic and any simple Lie algebra over the algebraically closed field Φ is isomorphic to one of these algebras or to one of five other Lie algebras. The orders of these exceptional Lie algebras are 14, 52, 78, 133 and 248.

3. **Construction of simple Lie algebras.** Let $\mathfrak{A}$ denote a simple associative algebra over Φ, an arbitrary field of characteristic 0, and n the degree of $\mathfrak{A}$ over its centrum.

Theorem 1. *If $\mathfrak{A}$ has centrum Σ and $n > 1$, then the derived algebra $\mathfrak{A}_l'$ is simple with Σ as its extended centrum.*[8]

$\mathfrak{A}_l'$ consists of sums of elements of the form $[a, b]$. For ξ in Σ, $[a, b]\xi = [a\xi, b] = [a, b\xi]$ and hence $\mathfrak{A}_l'$ as well as $\mathfrak{A}$ may be regarded as an algebra over Σ. $\mathfrak{A}$ is normal over Σ, $(\mathfrak{A}$ over $\Sigma)_\Omega \cong \Omega_n$ if Ω is the algebraic closure of Σ. Hence $(\mathfrak{A}_l$ over $\Sigma)_\Omega \cong \Omega_{nl}$ and $(\mathfrak{A}_l'$ over $\Sigma)_\Omega \cong \Omega_{nl}'$ a simple Lie algebra. Thus $\mathfrak{A}_l'$ is normal simple over Σ and the theorem follows from the remarks of §1.

The isomorphism $(\mathfrak{A}$ over $\Sigma)_\Omega \cong \Omega_n$ defines a representation of $\mathfrak{A}$ by matrices in Ω_n. In this representation the matrices of $\mathfrak{A}_l'$ have trace 0 and there are $m = n^2 - 1$ of them linearly independent over Ω since $(\mathfrak{A}_l'$ over $\Sigma)_\Omega \cong \Omega_{nl}'$. Hence the order of $\mathfrak{A}_l'$ over Φ is km if the order of Σ over Φ is k. Since $\mathfrak{A}_l'$ is not a subring of $\mathfrak{A}$, the order of the enveloping ring $\mathfrak{R}$ of $\mathfrak{A}_l'$ in $\mathfrak{A}$ exceeds m (over Σ) and so $\mathfrak{R} = \mathfrak{A}$.[9] The characteristic polynomials of the matrices of $\mathfrak{A}_l'$ all belong to the ring $\Sigma[\lambda]$.[10]

[7] Cartan [1], Chapter 5. We have substituted the algebra of class C with $n = 4$ for the isomorphic algebra of class B with $n = 5$.

[8] Landherr [1].

[9] If $\mathfrak{A}_l'$ is a subring of $\mathfrak{A}$, then the matrices of trace 0 form a subring of Ω_n and this is clearly impossible.

[10] Deuring [1], pp. 50–52.

THEOREM 2. *Suppose $\mathfrak{A}$ has centrum Σ of order k over Φ and the i.a.a. S of first kind. If n is odd, the order of $\mathfrak{S}_S$ over Φ is $\frac{1}{2}kn(n-1)$ and $\mathfrak{S}_S$ is simple if $n > 1$. If n is even, this order is $\frac{1}{2}kn(n-1)$ or $\frac{1}{2}kn(n+1)$. In the former case $\mathfrak{S}_S$ is simple if $n > 4$ and in the latter, if $n > 1$. The extended centrum of $\mathfrak{S}_S$ is Σ.*

By definition $\xi^S = \xi$ for every ξ in Σ. Hence if $b \in \mathfrak{S}_S$, $b\xi \in \mathfrak{S}_S$ also and $\mathfrak{S}_S$ may be regarded as an algebra over Σ. S is an i.a.a. in $\mathfrak{A}$ over Σ and has a unique extension to $(\mathfrak{A} \text{ over } \Sigma)_\Omega \cong \Omega_n$. The set of S-skew elements of $(\mathfrak{A}$ over $\Sigma)_\Omega$ is $(\mathfrak{S}_S \text{ over } \Sigma)_\Omega$. On the other hand, if n is odd, this set may be represented by the isomorphism $(\mathfrak{A} \text{ over } \Sigma)_\Omega \cong \Omega_n$ as the set of skew symmetric matrices, and if n is even, this set is representable as the set of skew symmetric matrices or as the set of matrices A such that $Q^{-1}A'Q = -A$, Q as in (1). Thus if n is restricted as in the hypothesis, $(\mathfrak{S}_S \text{ over } \Sigma)_\Omega$ is in one of Cartan's classes B, C or D. Hence $\mathfrak{S}_S$ is simple with Σ as its extended centrum.

By the above proof we obtain a representation of $\mathfrak{A}$ by matrices in Ω_n such that the matrices of $\mathfrak{S}_S$ form a basis for the matrices A such that $A' = -A$ or the matrices A such that $Q^{-1}A'Q = -A$. In either case, since the enveloping ring of $(\mathfrak{S}_S \text{ over } \Sigma)_\Omega$ is Ω_n, the enveloping ring of $\mathfrak{S}_S$ in $\mathfrak{A}$ is $\mathfrak{A}$ itself.

THEOREM 3. *If $\mathfrak{A}$ has centrum $\Sigma(q) \neq \Sigma$, $q^2 = \mu$ in Σ, $n > 1$ and S is an i.a.a. of second kind such that Σ is the set of S-symmetric elements of $\Sigma(q)$, then $\mathfrak{S}_S'$ is simple with Σ as its extended centrum.*

The above argument shows that $\mathfrak{S}_S$ and $\mathfrak{S}_S'$ may be regarded as algebras over Σ. If $a_0, \cdots, a_m$ is a basis for $\mathfrak{S}_S$ over Σ, it is easily seen that a_i is a basis for $\mathfrak{A}$ over $\Sigma(q)$. Hence $m = n^2 - 1$. If Ω is the algebraic closure of Σ and contains $\Sigma(q)$, $(\mathfrak{S}_S \text{ over } \Sigma)_\Omega \cong \Omega_{nl}$, $(\mathfrak{S}_S' \text{ over } \Sigma)_\Omega \cong \Omega_{nl}'$. Thus $\mathfrak{S}_S'$ is simple and has Σ as its extended centrum.

The order of $\mathfrak{S}_S'$ over Σ is m (km over Φ). Since the enveloping ring $\mathfrak{R}$ of $\mathfrak{S}_S'$ in $\mathfrak{A}$ is an algebra over Σ, as the author has shown,[11] $\mathfrak{R} = \mathfrak{A}$ if $n > 2$. Since $(\mathfrak{A} \text{ over } \Sigma(q))_\Omega \cong \Omega_n$, $\mathfrak{A}$ has a representation by matrices in Ω_n such that the matrices of $\mathfrak{S}_S'$ form a basis over Ω of all the matrices of trace 0 in Ω_n. The automorphism G induced by S in $\Sigma(q)$ can be extended to an automorphism $\bar{G}$ in Ω and $\bar{G}$ may be extended to an automorphism $\bar{G}$ in Ω_n by the definition $A^{\bar{G}} = (\alpha_{ij})^{\bar{G}} = (\alpha_{ij}^{\bar{G}})$.[12] It is readily seen that the correspondence $\sum a_i\omega_i \to -\sum a_i\omega_i^{\bar{G}}$ is an anti-automorphism $\bar{S}$ in $(\mathfrak{A} \text{ over } \Sigma(q))_\Omega$, i.e., in Ω_n extending S and the correspondence $A \to (A^{\bar{S}\bar{G}^{-1}})'$ is an automorphism in Ω_n leaving the elements of Ω invariant. It follows that there is a matrix S in Ω_n such that $A^{\bar{S}} = S^{-1}(A')^{\bar{G}}S = S^{-1}(A^{\bar{G}})'S$.[13] The characteristic polynomials of the matrices of $\mathfrak{A}$ all belong to $\Sigma(q)[\lambda]$. Now suppose A is a matrix of $\mathfrak{S}_S'$ and its characteristic polynomial $\lambda^n - \alpha_1\lambda^{n-1} + \cdots + (-1)^n\alpha_n$ is in $\Sigma[\lambda]$. Since $A = -S^{-1}(A')^{\bar{G}}S$, $\alpha_{2i+1}^{G} = -\alpha_{2i+1} = 0$. If this holds for all A in $\mathfrak{S}_S'$, it will hold for all matrices of trace 0 in Ω_n, and this is impossible if $n > 2$.

[11] Jacobson [2], p. 182.

[12] See, for example, the Anhang by Baer and Hasse to Steinitz, *Theorie der algebraischen Körpern.*

[13] Deuring [1].

4. **Simple Lie algebras of types** A, B, C, **and** D. If $\mathfrak{L}$ is a simple Lie algebra, Σ its extended centrum and $(\mathfrak{L}$ over $\Sigma)_\Omega$ is an algebra in Cartan's class A, B, C or D, then we say that $\mathfrak{L}$ has respectively *type* A, B, C or D. The order of $\mathfrak{L}$ over Σ is $n^2 - 1$, $\frac{1}{2}n(n-1)$, or $\frac{1}{2}n(n+1)$, and we suppose from now on that for type A, $n > 1$; for type B, $n > 5$; for type C, $n > 2$; and for type D, $n > 6$. The algebras obtained by Theorems 1 and 3 have type A, those of Theorem 2 have type B, C or D. If n is restricted as indicated, none of these algebras of different type are isomorphic since the extensions $(\mathfrak{L}$ over $\Sigma)_\Omega$ are not isomorphic. If an algebra of type A has a representation in Ω_n such that the representing matrices have characteristic polynomials all in $\Sigma[\lambda]$, then $\mathfrak{L}$ has *type* $\mathrm{A_I}$, otherwise it has *type* $\mathrm{A_{II}}$. Thus the Lie algebras of Theorem 1 have type $\mathrm{A_I}$ and it can be shown[14] that those of Theorem 3 have type $\mathrm{A_{II}}$ if $n > 2$. Hence an algebra of Theorem 1 is isomorphic to no algebra of Theorem 3 if $n > 2$.

We suppose in the sequel that if $\mathfrak{L}$ has type D, $n > 8$ (in addition to the conditions on n noted above for the other types). With this restriction we have

THEOREM 4. *If $\mathfrak{L}$ is a simple Lie algebra of type* A, B, C *or* D, *and Σ is its extended centrum, $\mathfrak{L}$ is isomorphic to one of the Lie algebras of Theorem* 1, 2 *or* 3.

For the algebras of type $\mathrm{A_I}$ this has been proved by Landherr [1] and for types $\mathrm{A_{II}}$, B, C and D by the author [1] and [2]. The condition $(\mathfrak{L}$ over $\Sigma)_\Omega \cong$ to an algebra in Cartan's class A, B, C or D defines a representation of $\mathfrak{L}$ by matrices in Ω_n. The algebra $\mathfrak{A}$ may be taken as the enveloping ring of the matrices in this representation. If $\mathfrak{L}$ has type $\mathrm{A_{II}}$, B, C or D, the i.a.a. S may be defined as the correspondence $\sum A_{i_1} \cdots A_{i_r} \to \sum (-1)^r A_{i_r} \cdots A_{i_1}$, where the A's are matrices of $\mathfrak{L}$.

5. **Isomorphisms and automorphisms.** The following theorems give conditions for the isomorphism of Lie algebras obtained by a construction of Theorems 1, 2 or 3 (n restricted as in the last section).

THEOREM 5. *If $\mathfrak{A}$ and $\mathfrak{B}$ are as in Theorem* 1 *and G is an isomorphism between $\mathfrak{A}_l'$ and $\mathfrak{B}_l'$ over Φ, then G may be realized as either an isomorphism or the negative of an anti-isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ over Φ.*[15]

By Theorem 1, $\mathfrak{A}$ and $\mathfrak{B}$ have isomorphic centrums and may be regarded as normal simple associative algebras over the same field $\Sigma \supset \Phi$. If Ω is the algebraic closure of Σ, we have seen that $\mathfrak{A}$ has a representation in Ω_n such that the matrices of $\mathfrak{A}_l'$ are $\sum_1^m A_i\xi_i$, $m = n^2 - 1$, $\xi \in \Sigma$ and the matrices of Ω_{nl}' are $\sum_1^m A_i\omega_i$, $\omega \in \Omega$. Similarly $\mathfrak{B}$ is representable in Ω_n such that the matrices of $\mathfrak{B}_l'$ are $\sum_1^m B_i\xi_i$, B_i linearly independent over Ω, and we may suppose that $B_i = A_i^G$. The automorphism G induced in Σ by the isomorphism G may be extended to an automorphism $\bar{G}$ in Ω and $\bar{G}$ may be extended to an automorphism $\bar{G}$ of

[14] Jacobson [2], p. 183.
[15] Cf. Landherr [1].

Ω_n by the definition $A^{\bar{G}} = (\alpha_{ij})^{\bar{G}} = (\alpha_{ij}^{G})$. Suppose $[A_i, A_j] = \sum A_k\gamma_{kij}, \gamma \in \Sigma$. Then $[B_i, B_j] = \sum B_k\gamma_{kij}^{G}$ and $[A_i^{\bar{G}}, A_j^{\bar{G}}] = \sum A_k^{\bar{G}}\gamma_{kij}^{G}$. Clearly $A_i^{G} \in \Omega'_{nl}$ and are linearly independent over Ω. Hence the correspondence $\sum A_i^{\bar{G}}\omega_i \rightarrow \sum B_i\omega_i$ is an automorphism of Ω'_{nl} over Ω. It follows by Weinstein's theorem that there is a matrix G such that either $B_i = G^{-1}A_i^{\bar{G}}G$ or $B_i = -G^{-1}(A_i^{\bar{G}})'G = -G^{-1}(A_i')^{\bar{G}}G$. In the former case the correspondence $A \rightarrow G^{-1}A^{\bar{G}}G$ is an automorphism in Ω_n which reduces to the isomorphism G for the matrices of $\mathfrak{A}'_l$ and in the latter $A \rightarrow G^{-1}(A')^{\bar{G}}G = A^U$ is an anti-automorphism in Ω_n whose negative is G for the matrices of $\mathfrak{A}'_l$. Since the enveloping ring of $\mathfrak{A}'_l$ is $\mathfrak{A}$, the matrices of $\mathfrak{A}$ have the form $\sum A_{i_1} \cdots A_{i_p}\xi_{i_1} \cdots \xi_{i_p}$, $\xi \in \Sigma$, and hence G or U maps the matrices of $\mathfrak{A}$ on those of $\mathfrak{B}$ and leaves the elements of Φ unaltered.

COROLLARY. *If G is an automorphism of $\mathfrak{A}'_l$ over Φ, it may be realized either as a unique automorphism or the negative of a unique anti-automorphism of $\mathfrak{A}$ over Φ.*

We identify $\mathfrak{A}$ and $\mathfrak{B}$ in Theorem 5. The uniqueness of the automorphism or anti-automorphism is an immediate consequence of the fact that the enveloping ring of $\mathfrak{A}'_l$ is $\mathfrak{A}$.

If S and T are two anti-automorphisms of $\mathfrak{A}$ over Φ, then ST is an automorphism of $\mathfrak{A}$ over Φ. Hence if $\mathfrak{G}$ is the group of automorphisms of $\mathfrak{A}'_l$ over Φ and $\mathfrak{G}_0$ the subgroup of the automorphism induced by the automorphisms of $\mathfrak{A}$ over Φ, then $\mathfrak{G}_0$ is an invariant subgroup of index 1 or 2 in $\mathfrak{G}$.

If $n = 2$, $\mathfrak{G}_0 = \mathfrak{G}$. For in this case $\mathfrak{A}$ is $\cong$ either to Σ_2 or to $\mathfrak{D}$ the generalized quaternion algebra whose basis is $1, i, j, k$, where

$$\begin{gathered} i^2 = \alpha, \qquad j^2 = \beta, \qquad k^2 = -\alpha\beta; \\ ij = -ji = k, \qquad jk = -kj = -i\beta, \qquad ki = -ik = -j\alpha. \end{gathered} \tag{2}$$

As we noted above Σ'_2 consists of the matrices of trace 0, and it is easily seen that its elements may also be characterized as the S_0-skew elements of Σ_2, where S_0 is the anti-automorphism $A \rightarrow Q^{-1}A'Q$, Q as in (1). In $\mathfrak{D}$ the correspondence $a = \alpha_0 + i\alpha_1 + j\alpha_2 + k\alpha_3 \rightarrow \bar{a} = \alpha_0 - i\alpha_1 - j\alpha_2 - k\alpha_3 \equiv a^{S_0}$ is an anti-automorphism whose skew elements are $i\alpha_1 + j\alpha_2 + k\alpha_3$, i.e., the elements of $\mathfrak{D}'_l$. Thus in either case $a^{S_0} = -a$ for the elements of $\mathfrak{A}'_l$. Hence the automorphism determined by $-S_0$ and by the identity automorphism are identical for the elements of $\mathfrak{A}'_l$ and $\mathfrak{G} = \mathfrak{G}_0$.

Suppose conversely that $\mathfrak{G} = \mathfrak{G}_0$. Then either there exists no anti-automorphisms of $\mathfrak{A}$ over Φ or there is an anti-automorphism S_0 such that $a^{S_0} = -a$ for all a in $\mathfrak{A}'_l$. For any ξ in Σ, $(a\xi)^{S_0} = a^{S_0}\xi^{S_0} = -a\xi^{S_0} = -a\xi$. Hence $\xi^{S_0} = \xi$ and $\mathfrak{S}_{S_0}$ the set of S_0-skew elements is a vector space over Σ. If $\mathfrak{S}_{S_0} > \mathfrak{A}'_l$ we must have $\mathfrak{S}_{S_0} = \mathfrak{A}$ which is impossible since $1 \in' \mathfrak{S}_{S_0}$. Thus $\mathfrak{S}_{S_0} = \mathfrak{A}'_l$ and if $a \in \mathfrak{A}'_l$ so does $a^3, a^5, \cdots$. It follows that tr $(a^3) = 0$ in the representation of $\mathfrak{A}$ in Ω_n, i.e., tr $(\Sigma A_i\xi_i)^3 = 0$ for all ξ in Σ. Hence tr $(\sum A_i\omega_i)^3 = 0$ for all ω in Ω. This is impossible if $n > 2$, for then there exist matrices B in Ω'_{nl} such that tr $B^3 \neq 0$. We have therefore shown that if $n > 2$ and $\mathfrak{A}$ over Φ has anti-automorphisms, then $\mathfrak{G} \neq \mathfrak{G}_0$.

Let $\mathfrak{F}$ denote the subgroup of $\mathfrak{G}_0$ consisting of the automorphisms induced by inner automorphisms of $\mathfrak{A}$. It is readily seen that $\mathfrak{F}$ is invariant in $\mathfrak{G}$ and is isomorphic to $\mathfrak{U}/\Sigma^*$, where $\mathfrak{U}$ is the group of units (non-singular elements) in $\mathfrak{A}$ and Σ^* is the group of multiple 1ξ, where $\xi \neq 0$ in the centrum.

THEOREM 6. *If $\mathfrak{A}$ and $\mathfrak{B}$ are as in Theorem 2, S and T the corresponding i.a.a.'s of first kind and G is an isomorphism between $\mathfrak{S}_S$ and $\mathfrak{S}_T$ over Φ, then G may be realized as an isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ over Φ.*

By Theorem 2, $\mathfrak{A}$ and $\mathfrak{B}$ may be regarded as normal simple over the same field Σ and by §4, $\mathfrak{S}_S$ and $\mathfrak{S}_T$ have the same type. We have seen that $\mathfrak{A}$ has a representation in Ω_n such that the matrices of $\mathfrak{S}_S$ are $\sum_1^m A_i\xi_i$, $m = \frac{1}{2}n(n-1)$ or $\frac{1}{2}n(n+1)$ and accordingly the A_i form a basis for either all skew-symmetric matrices (types B, D) or the matrices A such that $Q^{-1}A'Q = -A$, Q as in (1) (type C). The same result holds for $\mathfrak{B}$ and $\mathfrak{S}_T$ and we may suppose that $B_i = A_i^G$. As in the proof of Theorem 5 we obtain automorphisms $\bar{G}$ and $\bar{G}$ such that $A^{\bar{G}} = (\alpha_{ij}^{\bar{G}})$ and these reduce to G for the elements of Σ. Since $A_i^{\bar{G}}$ are skew-symmetric or satisfy $Q^{-1}(A_i^{\bar{G}})'Q = -A_i^{\bar{G}}$ if the A_i do, the correspondence $\sum A_i^{\bar{G}}\omega_i \to \sum B_i\omega_i$ is an automorphism in a Lie algebra of class B, C or D. Hence by the result quoted in §2, there is a matrix G such that $B_i = G^{-1}A_i^{\bar{G}}G$ and the correspondence $A \to G^{-1}A^{\bar{G}}G$ is an automorphism in Ω_n reducing to G for the matrices of $\mathfrak{S}_S$. Since the enveloping ring of $\mathfrak{S}_S$ is $\mathfrak{A}$, $A \to G^{-1}A^{\bar{G}}G$ is an isomorphism between the matrices of $\mathfrak{A}$ and $\mathfrak{B}$ over Φ.

COROLLARY 1. *If $\mathfrak{A}$ is as in Theorem 2 and S and T are i.a.a.'s such that $\mathfrak{S}_S$ and $\mathfrak{S}_T$ are isomorphic over Φ, then S and T are cogredient.*

If we identify $\mathfrak{A}$ and $\mathfrak{B}$ in Theorem 6 we obtain an automorphism G in $\mathfrak{A}$ over Φ mapping $\mathfrak{S}_S$ into $\mathfrak{S}_T$. Thus T and $G^{-1}SG$ have the same effect on the elements of $\mathfrak{S}_T$. Since the enveloping ring of $\mathfrak{S}_T$ is $\mathfrak{A}$, T and $G^{-1}SG$ have the same effect in $\mathfrak{A}$, i.e., $T = G^{-1}SG$.

COROLLARY 2. *Any automorphism G of $\mathfrak{S}_S$ over Φ may be realized by a unique automorphism of $\mathfrak{A}$ over Φ commutative with S.*

By Theorem 6 we have an automorphism G of $\mathfrak{A}$ over Φ which reduces to G for the elements of $\mathfrak{S}_S$. If $a \in \mathfrak{S}_S$, $a^S = -a$ and hence G and S commute for the elements of $\mathfrak{S}_S$. Since the enveloping ring of $\mathfrak{S}_S$ is $\mathfrak{A}$, $GS = SG$ for all elements of $\mathfrak{A}$. The uniqueness of the extension G follows for the same reason.

This corollary shows that the group $\mathfrak{G}$ of automorphisms of $\mathfrak{S}_S$ over Φ is isomorphic to the group of the automorphisms of $\mathfrak{A}$ over Φ commutative with S. Let $\mathfrak{F}$ denote the set of automorphisms of $\mathfrak{S}_S$ determined by inner automorphisms of $\mathfrak{A}$. Then $\mathfrak{F}$ is invariant in $\mathfrak{G}$ and $\cong \mathfrak{U}_S/\Sigma^*$, where $\mathfrak{U}_S$ is the totality of *S-orthogonal* elements g, i.e., $gg^S = 1\gamma \neq 0$ in Σ and Σ^* is the set 1ξ, $\xi \neq 0$ in Σ.

THEOREM 7. *If $\mathfrak{A}$ and $\mathfrak{B}$ are as in Theorem 3 with $n > 2$, S and T the corresponding i.a.a.'s of second kind and G is an isomorphism between $\mathfrak{S}_S'$ and $\mathfrak{S}_T'$ over Φ, then $\mathfrak{G}$ may be realized as an isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ over Φ.*

By Theorem 3 the subfields of symmetric elements of the centrums of $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic, $\cong \Sigma$, and $\mathfrak{A}$ and $\mathfrak{B}$ may be regarded as simple algebras over Σ

having the centrums $P_1 = \Sigma(q_1)$, $q_1^2 = \mu_1 \epsilon \Sigma$ and $P_2 = \Sigma(q_2)$, $q_2^2 = \mu_2 \epsilon \Sigma$, respectively. We may embed P_1 and P_2 in the same algebraic closure Ω. Then $\mathfrak{A}$ is representable by matrices in Ω_n such that the matrices of $\mathfrak{S}'_S$ are $\sum_1^m A_i\xi_i$, $m = n^2 - 1$, $\xi \epsilon \Sigma$ and the matrices of Ω'_{ni} are $\sum A_i\omega_i$. Similarly, we represent $\mathfrak{B}$ by matrices in Ω_n such that the matrices of $\mathfrak{S}'_T$ are $\sum B_i\xi_i$, where $B_i = A_i^G$. By the proof of Theorem 5 we obtain either an isomorphism G mapping $\mathfrak{A}$ over Φ into $\mathfrak{B}$ over Φ and reducing to G for the elements of $\mathfrak{S}'_S$, or we obtain an anti-isomorphism U such that $a^U = -a^G$ for the elements of $\mathfrak{S}'_S$. But in this case $UT = G$ is an isomorphism G of $\mathfrak{A}$ over Φ into $\mathfrak{B}$ over Φ reducing to G for the elements of $\mathfrak{S}'_S$. It follows, of course, that P_1 and P_2 are isomorphic.

COROLLARY 1. *If $\mathfrak{A}$ is as in Theorem 3 and S and T are i.a.a.'s such that $\mathfrak{S}'_S$ and $\mathfrak{S}'_T$ are isomorphic over Φ, then S and T are cogredient.*

COROLLARY 2. *Any automorphism G of $\mathfrak{S}'_S$ over Φ may be realized by a unique automorphism of $\mathfrak{A}$ over Φ commutative with S.*

The proofs are identical with those of Corollaries 1 and 2 to Theorem 6.

By the second of the present corollaries the group of automorphisms G of $\mathfrak{S}'_S$ over Φ is isomorphic to the subgroup of the automorphisms of $\mathfrak{A}$ over Φ commutative with S. The elements of $\mathfrak{G}$ determined by inner automorphisms for an invariant subgroup $\mathfrak{F} \cong \mathfrak{U}_S/P^*$, where $\mathfrak{U}_S$ is the totality of S-orthogonal elements and P^* is the set 1ξ, $\xi \neq 0$ in $P = \Sigma(q)$.

6. **Cogredience of i.a.a.'s and cogredience of matrices.** The above discussion reduces the problems of simple Lie algebras (with the exception of those having orders 14, 28, 52, 78, 133 and 248 over the extended centrum) to problems in simple associative algebras. In this section we shall see how the latter may be formulated as specific questions about matrices with elements in a division algebra.

It is well known that any simple associative algebra $\mathfrak{A}$ is isomorphic to a $\mathfrak{D}_r$, an r-rowed matrix algebra with elements in a division algebra $\mathfrak{D}$. $\mathfrak{D}$ is determined in the sense of isomorphism by $\mathfrak{A}$. We recall also that any automorphism leaving the elements of the centrum P invariant is inner.[16]

Suppose $\mathfrak{D}_r$ is involutorial anti-automorphic and S is an i.a.a. in $\mathfrak{D}_r$ over Φ. As has been shown by Albert[17] there exists an i.a.a. $a \to a^U$ in $\mathfrak{D}$ whose extension defined by $A^U = (a_{ij})^U = (a_{ji}^U)$ has the same effect in P as S and there exists a matrix S in $\mathfrak{D}_r$ such that $A^S = S^{-1}A^US$. The matrix S is determined to within multiplication by elements of P. Let Σ denote the symmetric part of P relative to U or S. If S is of first kind ($P = \Sigma$), then the matrix S is either U-symmetric ($S^U = S$) or U-skew ($S^U = -S$), and if S is of second kind, its matrix may be normalized to be U-symmetric and will then be determined to within multiplication by elements of Σ. Thus in any case we may associate with S a *ray* $\{S\}$ of non-singular U-symmetric or U-skew matrices consisting

[16] These results may be found in Deuring [1].

[17] Albert [1], p. 909.

of the multiples of the fixed U-symmetric or U-skew matrix S by the elements $\neq 0$ of Σ.

Let E_{ij} denote the matrix basis of $\mathfrak{D}_r$ such that $A = (a_{ij}) \equiv \sum a_{ij}E_{ij}$ and suppose G is an automorphism of $\mathfrak{D}_r$ over Φ. Let $E_{ij}^G = F_{ij}$. Then $F_{ij}F_{kl} = \delta_{jk}F_{il}$ and hence there is a matrix K such that $F_{ij} = K^{-1}E_{ij}K$. Hence the automorphism G' such that $A^{G'} = KA^GK^{-1}$ leaves the E_{ij} invariant. Since $\mathfrak{D}$ may be characterized as the totality of elements of $\mathfrak{D}_r$ commutative with the E_{ij}, G' maps $\mathfrak{D}$ into itself. Consider the i.a.a. $G'^{-1}UG'$. It maps E_{ij} into E_{ji} and $\mathfrak{D}$ into itself.

The inner automorphisms of $\mathfrak{D}_r$ form an invariant subgroup in the group of automorphisms of $\mathfrak{D}_r$ over Φ. The factor group determined is isomorphic to a subgroup of the Galois group of P over Φ. Let $G_1 = 1, G_2, \cdots, G_l$ be representatives of the cosets of this factor group. By the above we may suppose that the G_i leave the E_{ij} unaltered and map $\mathfrak{D}$ into itself. Hence the i.a.a. $U_q = G_q^{-1}UG_q$ map E_{ij} into E_{ji} and $\mathfrak{D}$ into itself. Now suppose $T = G^{-1}SG$ is cogredient to S. The automorphism G has the form G_qH, where $A^H = H^{-1}AH$. Then a simple computation yields $A^T = T^{-1}A^{U_q}T$, where $T = H^{U_q}S^{G_q}H$. We say that the rays $\{S_1\}$ and $\{S_2\}$ are *U-cogredient* if there exists a non-singular matrix H and an element $\rho \neq 0$ in Σ such that $S_2 = H^US_1H\rho$. Thus the i.a.a.'s cogredient to S have the form $A^T = T^{-1}A^{U_q}T$, where the rays $\{T\}$ and $\{S^{G_q}\}$ are U_q-cogredient for some q, and conversely. Similarly the condition that the inner automorphism G where $A^G = G^{-1}AG$ commute with S is that $G^USG = S\rho$, $\rho \neq 0$ in Σ.

7. Bilinear forms and cogredience of matrices. In this section we suppose $\mathfrak{D}$ is an arbitrary quasi-field[18] of characteristic $\neq 2$ and $a \to \bar{a}$ is a fixed i.a.a. in $\mathfrak{D}$. In particular if $\mathfrak{D}$ is commutative, $\mathfrak{D}$ may be the identity mapping. Let $\mathfrak{R}$ be a vector space of r $(< \infty)$ dimensions over $\mathfrak{D}$. A *bilinear form* f in $\mathfrak{R}$ is a function of pairs of vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathfrak{R}$ having values in $\mathfrak{D}$ such that

$$(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y}), \qquad (\mathbf{y}, \mathbf{x}_1 + \mathbf{x}_2) = (\mathbf{y}, \mathbf{x}_1) + (\mathbf{y}, \mathbf{x}_2),$$

$$(\mathbf{x}, \mathbf{y}a) = (\mathbf{x}, \mathbf{y})a, \qquad (\mathbf{x}a, \mathbf{y}) = \bar{a}(\mathbf{x}, \mathbf{y}), \qquad a \in \mathfrak{D}.$$

If $\mathbf{x}_1, \cdots, \mathbf{x}_r$ is a basis for $\mathfrak{R}$ and $(\mathbf{x}_i, \mathbf{x}_j) = s_{ij}$, then for $\mathbf{x} = \sum \mathbf{x}_i x_i$, $\mathbf{y} = \sum \mathbf{x}_i y_i$ we have $(\mathbf{x}, \mathbf{y}) = \sum \bar{x}_i s_{ij} y_j$, and conversely any matrix $S = (s_{ij})$ defines a bilinear form by this equation. S is called *the matrix* of f relative to the basis $\mathbf{x}_1, \cdots, \mathbf{x}_r$. A change to the basis $\mathbf{y}_1, \cdots, \mathbf{y}_r$, where $(\mathbf{y}_1, \cdots, \mathbf{y}_r) = (\mathbf{x}_1, \cdots, \mathbf{x}_r)H$, transforms S into the cogredient matrix $\bar{H}'SH$.

f is *Hermitian* if $\overline{(\mathbf{x}, \mathbf{y})} = (\mathbf{y}, \mathbf{x})$, *skew-Hermitian* if $\overline{(\mathbf{x}, \mathbf{y})} = -(\mathbf{y}, \mathbf{x})$. The conditions on S are respectively $\bar{S}' = S$ and $\bar{S}' = -S$. An element b of $\mathfrak{D}$ is represented by f if there exists a vector $\mathbf{u} \neq 0$ such that $(\mathbf{u}, \mathbf{u}) = b$. If f is Hermitian (skew-Hermitian), the elements represented by it are Hermitian (skew-Hermitian).

[18] I.e., not necessarily commutative field.

We suppose from now on that f is either Hermitian or skew-Hermitian. The vectors $\mathbf{u}$, $\mathbf{v}$ are *orthogonal* if $(\mathbf{u}, \mathbf{v}) = 0$ (or $(\mathbf{v}, \mathbf{u}) = 0$). If $\mathfrak{S}$ is a subspace of $\mathfrak{R}$, the totality of vectors orthogonal to all of the vectors of $\mathfrak{S}$ form a subspace $\mathfrak{S}'$ called the *orthogonal complement* of $\mathfrak{S}$. If $\mathfrak{R}' = 0$, f is said to be *degenerate*, otherwise *non-degenerate*. The conditions that $\mathbf{z} = \sum \mathbf{x}_i z_i$ belong to $\mathfrak{R}'$ are

$$(\mathbf{x}_i, \mathbf{z}) = \sum_{j=1}^{r} s_{ij} z_j = 0 \qquad (i = 1, \cdots, r).$$

Hence if the rank of S is s,[19] the dimensionality of $\mathfrak{R}'$ is $r - s$. In particular f is non-degenerate if and only if S is non-singular.

If $\mathfrak{D}$ is commutative, $\bar{d} = d$ and f is skew, then it is well known that f has a matrix of the form

$$\left(\begin{array}{cc|c} 0 & 1_\nu & \\ -1_\nu & 0 & \\ \hline & & 0 \end{array}\right), \qquad s = 2\nu. \tag{3}$$

Hence the rank of any skew matrix is even and the matrix is cogredient to (3). Any two skew matrices of the same rank are cogredient. In all other cases we proceed to show that f has a matrix of the form

$$\begin{pmatrix} b_1 & & & & & \\ & \ddots & & & & \\ & & b_s & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}. \tag{4}$$

LEMMA. *If $f \neq 0$, i.e., there exist vectors $\mathbf{u}$, $\mathbf{v}$ such that $(\mathbf{u}, \mathbf{v}) \neq 0$, then there exists a vector $\mathbf{u}$ such that $(\mathbf{u}, \mathbf{u}) \neq 0$.*

If $(\mathbf{u}, \mathbf{u}) = 0$ for all $\mathbf{u}$ in $\mathfrak{R}$ we have

$$(\mathbf{u}, \mathbf{v}) + (\mathbf{v}, \mathbf{u}) = (\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) - (\mathbf{u}, \mathbf{u}) - (\mathbf{v}, \mathbf{v}) = 0.$$

Hence for any a in $\mathfrak{D}$,

$$(\mathbf{u}, \mathbf{v}a) + (\mathbf{v}a, \mathbf{u}) = (\mathbf{u}, \mathbf{v})a + \bar{a}(\mathbf{v}, \mathbf{u}) = (\mathbf{u}, \mathbf{v})a - \bar{a}(\mathbf{u}, \mathbf{v}) = 0.$$

If f is Hermitian, set $a = \overline{(\mathbf{u}, \mathbf{v})} = (\mathbf{v}, \mathbf{u})$, $\bar{a} = (\mathbf{u}, \mathbf{v})$ in the second member of this equation, and we obtain $2(\mathbf{u}, \mathbf{v})(\mathbf{v}, \mathbf{u}) = 0$ contrary to the assumptions $f \neq 0$, $\mathfrak{D}$ of characteristic $\neq 2$. If f is skew-Hermitian, we obtain from the third member of the same equation that $\bar{a}(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{v})a$. Choose $\mathbf{u}$, $\mathbf{v}$ such that $(\mathbf{u}, \mathbf{v}) \neq 0$; then $\bar{a} = (\mathbf{u}, \mathbf{v})a(\mathbf{u}, \mathbf{v})^{-1}$. Thus the correspondence $a \to \bar{a}$ is an automorphism as well as an anti-automorphism. It follows that $\mathfrak{D}$ is commutative, $\bar{a} = a$, and we have the case previously treated and outside of the present consideration.

[19] Cf. van der Waerden, *Moderne Algebra*, I, 2d ed., p. 109, or II, 1st ed., p. 116.

If $f \neq 0$, let $\mathbf{u}_1$ be any vector such that $(\mathbf{u}_1, \mathbf{u}_1) = b_1 \neq 0$, and suppose we have already found k vectors $\mathbf{u}_1, \cdots, \mathbf{u}_k$ such that $(\mathbf{u}_i, \mathbf{u}_j) = 0$ if $i \neq j$ and $(\mathbf{u}_i, \mathbf{u}_i) = b_i \neq 0$. Let $\mathfrak{R}_k$ denote the space generated by $\mathbf{u}_1, \cdots, \mathbf{u}_k$ and $\mathbf{E}_k$ the transformation in $\mathfrak{R}$ defined by

$$\mathbf{x}\mathbf{E}_k = \sum_{i=1}^{k} \mathbf{u}_i(\mathbf{u}_i, \mathbf{u}_i)^{-1}(\mathbf{u}_i, \mathbf{x}).$$

$\mathbf{E}_k$ is a *projection* of $\mathfrak{R}$ on $\mathfrak{R}_k$, i.e., a linear transformation mapping $\mathfrak{R}$ on $\mathfrak{R}_k$, leaving the elements of $\mathfrak{R}_k$ invariant and such that $(\mathbf{x}, \mathbf{y}\mathbf{E}_k) = (\mathbf{x}\mathbf{E}_k, \mathbf{y})$ for any $\mathbf{x}, \mathbf{y}$.[20] Hence if we set $\mathbf{x} = \mathbf{x}\mathbf{E}_k + \mathbf{x}(\mathbf{1} - \mathbf{E}_k) = \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_1 = \mathbf{x}\mathbf{E}_k$, $\mathbf{x}_2 = \mathbf{x}(\mathbf{1} - \mathbf{E}_k)$, $\mathbf{x}_1$ will belong to $\mathfrak{R}_k$, $\mathbf{x}_2$ to $\mathfrak{R}_k'$ and $\mathfrak{R} = \mathfrak{R}_k \oplus \mathfrak{R}_k'$. If $f \neq 0$ in $\mathfrak{R}_k'$, there exists a vector $\mathbf{v}_{k+1}$ such that for $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}(\mathbf{1} - \mathbf{E}_k)$ we have $(\mathbf{u}_{k+1}, \mathbf{u}_{k+1}) = b_{k+1} \neq 0$. We repeat this process with $\mathfrak{R}_{k+1}$, the space of $\mathbf{u}_1, \cdots, \mathbf{u}_{k+1}$ and continue until, say for $\mathfrak{R}_s$, we obtain $f = 0$ in $\mathfrak{R}_s'$. We then choose $\mathbf{v}_{s+1}, \cdots, \mathbf{v}_r$ in $\mathfrak{R}$ such that $\mathbf{u}_{s+1} = \mathbf{v}_{s+1}(\mathbf{1} - \mathbf{E}_s), \cdots, \mathbf{u}_r = \mathbf{v}_r(1 - \mathbf{E}_s)$ form a basis for $\mathfrak{R}_s'$. The matrix of f relative to the basis $\mathbf{u}_1, \cdots, \mathbf{u}_r$ has the required diagonal form (4).

THEOREM 8. *Unless $\mathfrak{D}$ is commutative, $\bar{a} \equiv a$ and the matrix is skew, any Hermitian or skew-Hermitian matrix S is cogredient to a diagonal matrix.*

We note that b_i is any element represented by the form associated with S and any b_i may be replaced by $\bar{a}b_i a$, $a \neq 0$, $\bar{b}_i = \pm b_i$ according as S is Hermitian or skew-Hermitian. s in (4) is evidently the rank of S.

Let P denote the centrum of $\mathfrak{D}$ and Σ the subfield of symmetric elements of P. Then either $\mathrm{P} = \Sigma$ (first kind) or $\mathrm{P} = \Sigma(q)$, $q^2 = \mu$ in Σ (second kind). If $\mathfrak{D}$ has a finite basis over P, so has $\mathfrak{D}_r$, and we have seen that this algebra has a representation $A \to A_1$ by matrices in Ω_n, Ω the algebraic closure of P and n^2 is the order of $\mathfrak{D}_r$ over P. We write $N(A) = \det A_1$ and recall that $N(A) \in \mathrm{P}$. Suppose the i.a.a. is of first kind, $\mathrm{P} = \Sigma$. We have seen that $\bar{A}' \to S^{-1}A_1'S$, where S is either symmetric or skew in Ω_n. Hence if $B = \bar{G}'AG$, $N(B) = N(A)\gamma^2$, where $\gamma = N(G)$. Thus if Σ^* is defined as above, Σ_2^* is the subgroup of squares in Σ^* and $\Lambda = \Sigma^*/\Sigma_2^*$, then with every non-degenerate bilinear form f (or class of non-singular cogredient matrices) there is associated an element δ of Λ, namely, the element of Λ determined by $N(A)$ if A is a matrix of f. δ is called the *discriminant* of f (or of the class of matrices). If $\mathfrak{D}$ is of second kind, we have seen that if $A \to A_1$, then $\bar{A}' \to S^{-1}\bar{A}_1'S$ if $\omega \to \bar{\omega}$ is an extension in Ω of the automorphism induced by the i.a.a. in P. Thus $N(\bar{A}') = \overline{N(A)}$ and if A is Hermitian,[21] $N(A)$ is in Σ and if $B = \bar{G}'AG$, $N(B) = N(A)\gamma\bar{\gamma}$ and $\gamma = N(G)$. In this case we let Σ_N^* denote the elements of Σ^* which are norms $(\gamma\bar{\gamma})$ of elements of P and let $\Lambda = \Sigma^*/\Sigma_N^*$. Hence if f is a non-degenerate Hermitian form, we define its discriminant δ to be the element of Λ determined by $N(A)$ for A a matrix of f.

[20] $\mathbf{E}_k$ is uniquely determined by these properties. $\mathbf{1} - \mathbf{E}_k$ is a projection on $\mathfrak{R}_k'$.

[21] It is not necessary to consider skew-Hermitian matrices in this case. For if q is skew in P and A is skew-Hermitian, Aq is Hermitian.

Consider the special case where $\mathfrak{D}$ is the quaternion algebra with basis $1, i, j, k$ over Σ and products as in (2), and let the i.a.a. be defined by $\bar{a} = \alpha_0 - i\alpha_1 - j\alpha_2 - k\alpha_3$ if $a = \alpha_0 + i\alpha_1 + j\alpha_2 + k\alpha_3$. We may represent $\mathfrak{D}$ in Ω_2 by

$$(5)\quad 1 \to \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad i \to \begin{pmatrix} \sqrt{\alpha} & \\ & -\sqrt{\alpha} \end{pmatrix}, \quad j \to \begin{pmatrix} 0 & 1 \\ \beta & 0 \end{pmatrix}, \quad k \to \begin{pmatrix} 0 & \sqrt{\alpha} \\ -\beta\sqrt{\alpha} & 0 \end{pmatrix},$$

and $\mathfrak{D}_r$ is represented in Ω_{2r} by replacing the element a_{ij} in A by its corresponding two-rowed matrix in (5). The condition $\bar{a} = a$ is equivalent to $a \in \Phi$ and $\bar{a} = -a$ is equivalent to $\operatorname{tr} a = a + \bar{a} = 0$. Then $a^2 + N(a) = 0$, where $N(a)$ is defined as above and $= a\bar{a} = \bar{a}a = -\alpha\alpha_1^2 - \beta\alpha_2^2 + \alpha\beta\alpha_3^2$. If $\bar{A}' = A$, we have seen that A is cogredient to B given by (4) and $\bar{b}_i \in \Sigma$. If A is non-singular, we have $N(B) \in \Sigma_2^*$ and hence $\delta(A) = 1$.[22] If $\bar{A}' = -A$ is non-singular, then $\delta(A)$ is the element of Λ determined by $N(b_1 \cdots b_r) = N(b_1) \cdots N(b_r)$.

8. **Real closed case**. Suppose Φ is real closed.[23] The division algebras over Φ are Φ, $\mathrm{P} = \Phi(\sqrt{-1})$ and $\mathfrak{D}$, Hamilton's quaternion algebra. Hence the simple associative algebras over Φ are Φ_n, P_n and $\mathfrak{D}_r$. Φ_n has the i.a.a. $A \to A'$ of first kind; P_n has the i.a.a. $A \to A'$ of the first kind and the i.a.a. $A \to \bar{A}'$, where $\bar{a}$ is the conjugate of a in P; $\mathfrak{D}_r$ has the i.a.a. $A \to \bar{A}'$, where $\bar{a}$ is the conjugate quaternion of a. The automorphisms of Φ_n and $\mathfrak{D}_r$ are all inner, but P_n has the outer automorphism $A \to \bar{A}$.

(a) Suppose $S \in \Phi_n$ and $S' = -S$ is non-singular. Then S is cogredient to Q given by (1). If $S' = S$ is non-singular, by Theorem 8 S is cogredient to one of the matrices

$$(6)\qquad S_p = \begin{pmatrix} \left.\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}\right\} p & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{pmatrix} \qquad (p = 0, 1, \cdots, n).$$

By Sylvester's theorem on the invariance of the signature $2p - n$, distinct S_p's are not cogredient. Clearly no H exists such that $Q = H'S_pH\rho$ since S_p is symmetric and Q is skew. Suppose now that $S_q = H'S_pH\rho$. If $\rho > 0$, we have $S_q = K'S_pK$ for $K = H\sqrt{\rho}$, and if $\rho < 0$, then $S_q = -K'S_pK$ for $K = H\sqrt{-\rho}$. Hence either $q = p$ or $q = n - p$. It follows from the above theory that any i.a.a. in Φ_n is cogredient either to Q: $A \to Q^{-1}A'Q$ or to one of the

[22] In this case the discriminant as defined above does not distinguish between any non-singular Hermitian matrices. However, a determinant can be defined for these matrices having the desired properties. See Moore [1], Chapter II.

[23] For the classification of Lie algebras over Φ, cf. Cartan [2].

$[\frac{1}{2}n] + 1$ i.a.a.'s S_p: $A \to S^{-1}A'S_p$, where $p = 0, 1, \cdots, [\frac{1}{2}n]$ and none of these i.a.a.'s is cogredient.

(b) Let $S \epsilon P_n$ and $S' = -S$ be non-singular. S is cogredient to Q in (1). Any non-singular $S' = S$ is cogredient to 1, the identity matrix. It follows that any i.a.a. in P_n is cogredient either to Q: $A \to Q^{-1}A'Q$ or to U: $A \to A'$ and these two are not cogredient.

Suppose $S \epsilon P_n$ is non-singular and $\bar{S}' = S$. S is cogredient to S_p in (6) and by Sylvester's theorem distinct S_p's are not cogredient. Hence any i.a.a. of second kind in P_n is cogredient to S_p: $A \to S_p^{-1}\bar{A}'S_p$, where $p = 0, 1, \cdots, [\frac{1}{2}n]$. The condition that S_p and S_q be cogredient i.a.a.'s is that $S_q = \bar{H}'S_pH\rho$ or $S_q = \bar{H}'\bar{S}_pH\rho = \bar{H}'S_pH\rho$. Hence as in (a) no two of the S_p are cogredient for $p = 0, \cdots, [\frac{1}{2}n]$.

(c) Let $S \epsilon \mathfrak{D}_r$ and $\bar{S}' = -S$ be non-singular. Then S is cogredient to

$$V = \begin{pmatrix} v & & & \\ & v & & \\ & & \ddots & \\ & & & v \end{pmatrix}, \tag{7}$$

where v is any element in $\mathfrak{D}$ such that $\bar{v} = -v$. By Theorem 8 it suffices to prove the

Lemma. *If $\bar{v} = -v$, $\bar{u} = -u \epsilon \mathfrak{D}$, then v and u are cogredient in $\mathfrak{D}$.*

Note first that v and $v\alpha$ are cogredient if $\alpha > 0$ in Φ. For $\Phi(v) \cong \Phi(\sqrt{-1})$ and hence $\alpha = \bar{a}a$, a in $\Phi(v)$, and $v\alpha = \bar{a}va$. The condition $\bar{v} = -v$ is equivalent to $\text{tr}\,(v) = v + \bar{v} = 0$. If $\alpha = [N(u)/N(v)]^{\frac{1}{2}}$ $(N(u) = u\bar{u} = \bar{u}u)$, then $N(v\alpha) = N(u)$. Since $\text{tr}\,(v\alpha) = \text{tr}\,(u) = 0$, $v\alpha$ and u are similar, say $v\alpha = g^{-1}ug = \gamma^{-1}\bar{g}ug$, $\gamma = N(g)$. Hence $v\alpha\gamma$ and u are cogredient and v and u are cogredient also.

If $\bar{S}' = S$ is non-singular in $\mathfrak{D}_r$, it is cogredient to S_p given by (5). Sylvester's theorem holds in this case also[24] and the argument above shows that any i.a.a. in $\mathfrak{D}_r$ is cogredient to V: $A \to V^{-1}\bar{A}'V$, V as in (7), or to S_p: $A \to S_p^{-1}\bar{A}'S_p$, $p = 0, 1, \cdots, [\frac{1}{2}r]$, and no two of these are cogredient.

If Φ is extended to its algebraic closure P, Φ_n extends to P_n and the i.a.a. Q extends to Q in P_n. Hence the Lie algebra $\mathfrak{S}_Q$ in Φ_n is of type C if $n > 1$. The i.a.a. S_p in Φ_n is cogredient in P_n to U and hence the Lie algebras $\mathfrak{S}_{S_p}$ in Φ_n have type B or D according as n is odd and > 1 or even and > 4.

The extension $\mathfrak{D}_{rP} \cong P_{2r}$. The matrices of $\mathfrak{S}_V$ in $\mathfrak{D}_r$ are $A = (a_{ij})$, where $a_{ji} = -v^{-1}\bar{a}_{ij}v$. In particular for $i = j$ we have $a_{ii} = -v^{-1}\bar{a}_{ii}v$ and hence $\overline{va}_{ii} = va_{ii} \epsilon \Phi$ and $a_{ii} = v\alpha_{ii}$, $\alpha_{ii} \epsilon \Phi$. It follows that the order of $\mathfrak{S}_V$ over Φ is $r + 4r(r-1)/2 = 2r(2r-1)/2$.[25] Thus $\mathfrak{S}_{VP} \cong \mathfrak{S}_U$ in P_{2r} and $\mathfrak{S}_V$ has type D if $r > 2$. Similarly we obtain the order of $\mathfrak{S}_{S_p}$ in $\mathfrak{D}_r$ as $3r + 4r(r-1)/2 = 2r(2r+1)/2$ and hence if $r > 0$, $\mathfrak{S}_{S_p}$ has type C. These results give the follow-

[24] Moore [1], p. 193.

[25] This can be seen more directly by examining the representation of $\mathfrak{D}_r$ given in §7.

ing list of non-isomorphic simple Lie algebras over Φ together with their automorphisms.

A_{I}. The algebras Φ_n', P_n' and $\mathfrak{D}_r'$ for $n, r > 1$. The orders over Φ are respectively $n^2 - 1$, $2(n^2 - 1)$ and $(2r)^2 - 1$. As before let $\mathfrak{G}$ be the group of automorphisms, $\mathfrak{G}_0$ the subgroup of the automorphisms determined by automorphisms of the associative algebra and $\mathfrak{F}$ the subgroup of $\mathfrak{G}_0$ of elements given by inner automorphisms of the associative algebra. For Φ_2' and $\mathfrak{D}_2'$ we have $\mathfrak{G} = \mathfrak{G}_0 = \mathfrak{F} \cong \mathfrak{U}/\Phi^*$, $\mathfrak{U}$ the group of units and Φ^* the elements $\neq 0$ of Φ. For P_2' we have $\mathfrak{G} = \mathfrak{G}_0 > \mathfrak{F} \cong \mathfrak{U}/\mathrm{P}^*$ and $\mathfrak{F}$ has index 2 in $\mathfrak{G}$. An element of $\mathfrak{G}$ not in $\mathfrak{F}$ is $A \to \bar{A}$. If $r, n > 2$ we see for Φ_n' and $\mathfrak{D}_r'$ that $\mathfrak{G} > \mathfrak{G}_0 = \mathfrak{F} \cong \mathfrak{U}/\Phi^*$. $\mathfrak{G}_0$ has index 2 in $\mathfrak{G}$ and an element of $\mathfrak{G}$ not in $\mathfrak{G}_0$ is $A \to A'$ (in Φ_n') or $A \to \bar{A}'$ (in $\mathfrak{D}_r'$). For P_n', $n > 2$, we have $\mathfrak{G} > \mathfrak{G}_0 > \mathfrak{F} \cong \mathfrak{U}/\mathrm{P}^*$. An element of $\mathfrak{G}$ not in $\mathfrak{G}_0$ is $A \to A'$ and one in $\mathfrak{G}_0$ not in $\mathfrak{F}$ is $A \to \bar{A}$. These automorphisms commute and hence $\mathfrak{G}/\mathfrak{F}$ is isomorphic to the Vierergruppe.

A_{II}. These are the algebras $\mathfrak{S}_{S_p}$, $p = 0, \cdots, [\frac{1}{2}n]$, in P_n, $n > 2$. The orders are $n^2 - 1$ over Φ. The automorphism $A \to \bar{A}$ commutes with S_p. Hence $\mathfrak{F}$ has index 2 in $\mathfrak{G}$. $\mathfrak{F} \cong \mathfrak{U}_{S_p}/\mathrm{P}^*$, where $\mathfrak{U}_{S_p}$ consists of the matrices U such that $\bar{U}'S_pU = S_p\rho$, where S_p is the matrix (6) associated with the i.a.a. S_p. Comparing signature we obtain that $\rho > 0$ unless n is even and $p = \frac{1}{2}n$. In the latter case S_p and $-S_p$ are cogredient matrices and hence there are U's in $\mathfrak{U}_{S_p}$ for which $\rho < 0$. The totality of U's with $\rho > 0$ form an invariant subgroup $\mathfrak{U}_{S_p}^+$ of index 2 in $\mathfrak{U}_{S_p}$ containing P^*. Hence $\mathfrak{F} \cong \mathfrak{U}_{S_p}/\mathrm{P}^*$ has an invariant subgroup $\mathfrak{F}^+$ of index 2. $\mathfrak{F}^+$ is invariant in $\mathfrak{G}$ also and $\mathfrak{G}/\mathfrak{F}^+ \cong$ the Vierergruppe.

B. The algebras $\mathfrak{S}_{S_p}$ in Φ_n and $\mathfrak{S}_U$ in P_n for n odd and > 5 and $p = 0, \cdots, [\frac{1}{2}n]$. The orders are respectively $n(n-1)/2$ and $2n(n-1)/2$. For $\mathfrak{S}_{S_p}$ we have $\mathfrak{G} = \mathfrak{F} \cong \mathfrak{U}_{S_p}/\Phi^*$, where $\mathfrak{U}_{S_p}$ is the set of matrices U such that $U'S_pU = S_p\rho$. ρ is necessarily > 0. For $\mathfrak{S}_U$, $\mathfrak{F}$ has index 2 in $\mathfrak{G}$ since $A \to \bar{A}$ is in $\mathfrak{G}$ but not in $\mathfrak{F}$. In this case $\mathfrak{F} \cong \mathfrak{U}_U/\mathrm{P}^*$, where $\mathfrak{U}_U$ is the group of matrices U such that $U'U = 1\rho$, ρ in P.

C. The algebras $\mathfrak{S}_Q$ in Φ_n and $\mathfrak{S}_Q$ in P_n for n even and > 2 and the algebras $\mathfrak{S}_{S_p}$ in $\mathfrak{D}_r$ for $r > 1$. The orders are respectively $n(n+1)/2$, $2n(n+1)/2$ and $2r(2r+1)/2$. For $\mathfrak{S}_Q$ in Φ_n, $\mathfrak{G} = \mathfrak{F} \cong \mathfrak{U}_Q/\Phi^*$, $\mathfrak{U}_Q$ the set of matrices U such that $U'QU = Q\rho$. The matrices U such that $\rho > 0$ form an invariant subgroup $\mathfrak{U}_Q^+$ of index 2 in $\mathfrak{U}_Q$ and $\mathfrak{U}_Q^+$ contains Φ^*. Hence $\mathfrak{F}$ has an invariant subgroup $\mathfrak{F}^+$ of index 2. For $\mathfrak{S}_Q$ in P_n, $\mathfrak{F}$ has index 2 in $\mathfrak{G}$ and is isomorphic to $\mathfrak{U}_Q/\mathrm{P}^*$, $\mathfrak{U}_Q$ the matrices U such that $U'QU = Q\rho$, ρ in P. For $\mathfrak{S}_{S_p}$ in $\mathfrak{D}_r$, $\mathfrak{G} = \mathfrak{F} \cong \mathfrak{U}_{S_p}/\Phi^*$. If $U \in \mathfrak{U}_{S_p}$, i.e., $\bar{U}'S_pU = S_p\rho$, then $\rho > 0$ unless r is even and $p = \frac{1}{2}r$ and in this case the U's with $\rho > 0$ form an invariant subgroup $\mathfrak{U}_{S_p}^+$ of index 2 in $\mathfrak{U}_{S_p}$ and containing Φ^*. Hence for $p = \frac{1}{2}r$, $\mathfrak{F}$ has an invariant subgroup of index 2.

D. The algebras $\mathfrak{S}_{S_p}$ in Φ_n and $\mathfrak{S}_U$ in P_n for n even and > 8 and $p = 0, \cdots, \frac{1}{2}n$ and the algebras $\mathfrak{S}_V$ in $\mathfrak{D}_r$ for $r > 4$. The orders are respectively $n(n-1)/2$, $2n(n-1)/2$ and $2r(2r-1)/2$. For $\mathfrak{S}_{S_p}$, $\mathfrak{G} = \mathfrak{F} \cong \mathfrak{U}_{S_p}/\Phi^*$. If $p = \frac{1}{2}n$, $\mathfrak{F}$ has an invariant subgroup $\mathfrak{F}^+$ of index 2 determined as above. For $\mathfrak{S}_U$ in P_n, $\mathfrak{F}$ has

index 2 in $\mathfrak{G}$. The automorphism $A \to \bar{A}$ is in $\mathfrak{G}$ but not in $\mathfrak{F}$. $\mathfrak{F} \cong \mathfrak{S}_U/P^*$. For $\mathfrak{S}_V$ in $\mathfrak{D}_r$, $\mathfrak{G} = \mathfrak{F} \cong \mathfrak{U}_V/\Phi^*$. $\mathfrak{U}_V$ is the group of matrices U such that $\bar{U}'VU = V\rho$. As before $\mathfrak{F}$ has an invariant subgroup $\mathfrak{F}^+$ of index 2.

9. **$\mathfrak{p}$-adic fields**. The division algebras over a $\mathfrak{p}$-adic field Φ have been determined by Hasse.[26] However, their automorphisms seem not to have been treated in the general case of non-normal algebras. We shall therefore restrict our attention to normal simple Lie algebras $\mathfrak{L}$ and their automorphisms. Since any simple Lie algebra is normal simple over its extended centrum, the only loss of generality here is in connection with the automorphisms of $\mathfrak{L}$.

There are $\phi(m)$ (Euler function) distinct normal division algebras over m^2 over Φ. These are anti-automorphic if and only if $m = 1$ or 2. If $m > 2$ the $\phi(m)$ algebras may be paired into $\frac{1}{2}\phi(m)$ pairs consisting of an algebra and its anti-isomorphic algebra. Any normal simple algebra of order n^2 over Φ has the form $\mathfrak{D}_r$, where $\mathfrak{D}$ is a normal division algebra of order m^2 and $n = mr$. It follows that there are $\sum_{\substack{m/n \\ m>2}} \frac{1}{2}\phi(m) + 2 = [\frac{1}{2}(n + 3)]$ non-isomorphic normal simple Lie algebras of type A_I and order $n^2 - 1$ over Φ to which any Lie algebra of this type is isomorphic.

As the author has shown, the only division algebras of second kind over Φ are the commutative fields $P = \Phi(q)$, $q^2 = \mu$ in Φ.[27] As above we denote the set of non-zero elements of Φ by Φ^* the set of squares in Φ^* by Φ_2^* and Φ^*/Φ_2^* by Λ. If $\mathfrak{p} \nmid 2$ the order of Λ is 4 and if $\mathfrak{p} \mid 2$ its order is 2^{l+2}, where l is the order of Φ over the field of 2-adic numbers.[28] Hence for $\mathfrak{p} \nmid 2$ we have three non-isomorphic quadratic extensions of Φ and for $\mathfrak{p} \mid 2$, $2^{l+2} - 1$ such extensions. To obtain the Lie algebras of type A_{II} we consider the algebras P_n for $n > 2$ and classify the i.a.a.'s of second kind in P_n and hence the rays of Hermitian matrices. It has been shown by Landherr[29] that if n is odd there are two cogredience classes, and if n is even one class. In the respective cases we have for $\mathfrak{p} \nmid 2$ three or six non-isomorphic Lie algebras and for $\mathfrak{p} \mid 2$, $2^{l+2} - 1$ or $2(2^{l+2} - 1)$ such algebras.

To obtain the Lie algebras of type B, C and D we consider the anti-automorphic associative algebras of first kind Φ_n and $\mathfrak{D}_r$, $\mathfrak{D}$ the quaternion division algebra over Φ. The cogredience problem for rays of symmetric matrices has been treated by Hasse. We refer to his papers ([2], [3], and [4]) recalling merely that the number of cogredience classes depends only on $\mathfrak{p}$ and on the residue of n mod 2. The corresponding Lie algebras are of type B or D according as n is odd and > 1 or even and > 4. If we consider the i.a.a.'s associated with skew matrices, we obtain one other class, the resulting Lie algebra having type C if $n > 1$.

[26] Hasse [1].
[27] Jacobson [4].
[28] Hasse [2], p. 115.
[29] Landherr [1]. Landherr obtains here the algebras of type A_{II} by a method different from ours.

It remains to discuss $\mathfrak{D}_r$, where $\mathfrak{D}$ and the i.a.a. are as in §7. For $a = \alpha_0 + i\alpha_1 + j\alpha_2 + k\alpha_3$ we have $N(a) = \alpha_0^2 - \alpha\alpha_1^2 - \beta\alpha_2^2 + \alpha\beta\alpha_3^2$. Since any non-degenerate quadratic form in four variables over Φ represents all elements of Φ^*, it follows that every element of Φ^* is a norm of an element in $\mathfrak{D}$. Hence the elements $b_i = \bar{b}_i = \beta_i \in \Phi^*$ in (4) may be replaced by 1 and any non-singular Hermitian matrix in $\mathfrak{D}_r$ is cogredient to 1. There is therefore a single cogredience class of i.a.a.'s given by Hermitian matrices. A simple computation shows that the order of the corresponding Lie algebra $\mathfrak{S}_U$ is $2r(2r + 1)/2$ and hence $\mathfrak{S}_U$ has type C if $r > 0$.

Lemma 1. *The skew Hermitian elements u and v $(\neq 0)$ in $\mathfrak{D}$ are cogredient if and only if $\Phi(u)$ and $\Phi(v)$ are isomorphic, i.e., $N(u)/N(v) \in \Phi_2^*$.*

Clearly if $u = \bar{g}vg$ and $\mu = N(u)$, $\nu = N(v)$, $\gamma = N(g)$, then $\mu = \gamma^2\nu$. Since $u^2 = -\mu$, $v^2 = -\nu$, $\Phi(u) \cong \Phi(v)$. Now suppose $\Phi(u) = \Phi(v)$. Then $v = u\rho$. If $\rho = N(q)$ for q in $\Phi(u)$, we obtain $v = \bar{q}uq$ is cogredient to u. If ρ is not a norm, $\rho = \sigma\tau$, where σ is any non-norm and τ is a norm in $\Phi(u)$. There is an element w in $\mathfrak{D}$ such that $w^{-1}uw = -u$ and $w^2 = \sigma$ is not a norm in $\Phi(u)$. Thus $\bar{w}uw = u\sigma$ and $v = u\rho = u\sigma\tau$ is cogredient to $u\sigma$ and to u. If $\Phi(u)$ is isomorphic to $\Phi(v)$, $\mu = \gamma^2\nu$ and hence $N(u\gamma) = N(v)$. There exists an element z in $\mathfrak{D}$ such that $u\gamma = z^{-1}vz$. Therefore v is cogredient to $u\gamma\zeta$ if $\zeta = N(z)$ and hence v is cogredient to u.

We recall that $\mathfrak{D}$ contains all quadratic extensions P of Φ.[30] It follows that if $\mathfrak{p} \nmid 2$, there are 3, and if $\mathfrak{p} \mid 2$, $2^{l+2} - 1$ cogredience classes of skew-Hermitian elements of $\mathfrak{D}$.

Lemma 2. *If f is a skew-Hermitian form and* dim $\mathfrak{R}$ *over* $\mathfrak{D} = r > 3$, *then f represents* 0.

If f is degenerate, this is trivial. Hence suppose it non-degenerate and let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r$ be a basis for $\mathfrak{R}$ such that the matrix of f is (4) with $s = r$. Then for $\mathbf{w} = \sum \mathbf{u}_i w_i$ we have $(\mathbf{w}, \mathbf{w}) = \sum \bar{w}_i b_i w_i$. By the preceding lemma we may choose w_i so that $\bar{w}_i b_i w_i = b_i\beta_i$, where β_i is arbitrary in Φ. Since the space $\mathfrak{S}$ of skew-Hermitian elements of $\mathfrak{D}$ has three dimensions over Φ and $r > 3$, we may choose β_i not all 0 such that $\sum b_i\beta_i = \sum \bar{w}_i b_i w_i = 0$.

Lemma 3. *If f is a non-degenerate skew-Hermitian form and represents* 0, *then f represents every skew-Hermitian element of $\mathfrak{D}$.*

Suppose $(\mathbf{u}, \mathbf{u}) = 0$, $\mathbf{u} \neq 0$. Since f is non-degenerate, there is a vector $\mathbf{v}$ such that $(\mathbf{u}, \mathbf{v}) = d \neq 0$. Then replacing $\mathbf{v}$ by $\mathbf{v}d^{-1}$, we may suppose $(\mathbf{u}, \mathbf{v}) = 1$. Let $(\mathbf{v}, \mathbf{v}) = e$. If u is any element of $\mathfrak{S}$, $u - e \in \mathfrak{S}$, and it is easily seen that if $\mathbf{w} = \mathbf{u}\frac{1}{2}(e - u) + \mathbf{v}$, $(\mathbf{w}, \mathbf{w}) = u$.

Lemma 4. *If $r > 2$ and f is non-degenerate, then f represents every skew-Hermitian element of $\mathfrak{D}$.*

By Lemmas 2 and 3 we may suppose that $r = 3$ and that f does not represent 0. Then b_1, b_2, b_3 in Lemma 2 are linearly independent and hence form a basis for $\mathfrak{S}$. Hence if $u \in \mathfrak{S}$, $u = \sum b_i\beta_i = \sum \bar{w}_i b_i w_i$ is represented by f.

[30] Deuring [1], p. 113.

LEMMA 5. *If $r = 2$ and f and g are non-degenerate skew-Hermitian forms, then there is an element $u \neq 0$ represented by both f and g.*

By Lemma 3 we may suppose that neither f nor g represents 0. Relative to suitable bases of $\mathfrak{R}$ the matrices of f and g are respectively

$$B_0 = \begin{pmatrix} b_1 & \\ & b_2 \end{pmatrix}, \qquad C_0 = \begin{pmatrix} c_1 & \\ & c_2 \end{pmatrix}. \tag{8}$$

Consider the form in four-space over $\mathfrak{D}$ having the matrix

$$\begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & -c_1 & \\ & & & -c_2 \end{pmatrix}.$$

By Lemma 2 there are elements w_1, w_2, w_3 and w_4 not all 0 such that $\bar{w}_1 b_1 w_1 + \bar{w}_2 b_2 w_2 = \bar{w}_3 c_1 w_3 + \bar{w}_4 c_2 w_4 = u$, $u \neq 0$, since f does not represent 0.

LEMMA 6. *The matrices B_0 and C_0 in* (8) *are cogredient if and only if they have the same discriminant.*

By Lemma 5 and the proof of Theorem 8 these matrices are cogredient respectively to

$$\begin{pmatrix} u & \\ & b_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} u & \\ & c_3 \end{pmatrix}.$$

Since the discriminants are equal, $N(b_3) = \gamma^2 N(c_3)$. Hence by Lemma 1, b_3 and c_3 are cogredient and the above matrices are cogredient also.

If f is any non-degenerate skew-Hermitian form in $\mathfrak{R}$ over $\mathfrak{D}$ and $u \neq 0$ is arbitrary in $\mathfrak{D}$, it follows by the proof of Theorem 8 that we may suppose $b_1 = b_2 = \cdots = b_{r-2} = u$ in (4), i.e., any non-singular skew-Hermitian matrix is cogredient to a matrix of the form

$$\begin{bmatrix} u & & & & \\ & \ddots & & & \\ & & u & & \\ & & & b_{r-1} & \\ & & & & b_r \end{bmatrix}, \tag{9}$$

where $u \neq 0$ is arbitrary in $\mathfrak{S}$. Hence we have the following criterion:

THEOREM 9. *Two non-singular skew-Hermitian matrices with elements in a $\mathfrak{p}$-adic quaternion algebra are cogredient if and only if they have the same discriminant.*

We may suppose that B is the matrix (9) and C is diagonal with elements $u, \cdots, u, c_{r-1}, c_r$. If the discriminants $\delta(B) = \delta(C)$, then $\delta(B_0) = \delta(C_0)$ for

$$B_0 = \begin{pmatrix} b_{r-1} & \\ & b_r \end{pmatrix} \quad \text{and} \quad C_0 = \begin{pmatrix} c_{r-1} & \\ & c_r \end{pmatrix},$$

and hence by Lemma 6 these matrices are cogredient. But then B and C are cogredient.

Let γ be any element of Φ^* and c an element of $\mathfrak{D}$ such that $N(c) = \gamma$. $\Phi(c)$ is a quadratic field and contains v a skew-Hermitian element. Then $\Phi(c) = \Phi(v)$. As we have seen, there is an element w_1 in $\mathfrak{D}$ such that $w_1^{-1}vw_1 = -v$, and since $w_1^2 = \sigma$, w_1 is in $\mathfrak{S}$. $w_2 = w_1^{-1}c$ satisfies these conditions also. Hence $\gamma = N(w_1)N(w_2)$ and $w_1, w_2 \in \mathfrak{S}$. If we replace w_2 by w_2v^{-1} and set $w_3 = v$, then $w_1, w_2, w_3 \in \mathfrak{S}$ and $\gamma = N(w_1)N(w_2)N(w_3)$. Thus if r is even, the diagonal matrix with elements $u, \cdots, u, v_1, v_2$ has norm $= \gamma\delta^2$, and if r is odd and > 1, the diagonal matrix $u, \cdots, u, v_1, v_2, v_3$ has norm $= \gamma\delta^2$. In either case if $r > 1$ every element of Λ is a discriminant of a skew-Hermitian matrix. If $\mathfrak{p} \nmid 2$, there are four cogredience classes, and if $\mathfrak{p} \mid 2$, there are 2^{l+2} such classes. The corresponding Lie algebras have type 0 if $r > 2$ (order $= 2r(2r-1)/2$).

The automorphisms of the above Lie algebras may be discussed along the lines indicated in the last section.

Bibliography

A. A. Albert.

1. *Involutorial simple algebras and real Riemann matrices*, Annals of Math., vol. 36 (1935), pp. 886–964.

E. Cartan.

1. *Thèse*, Paris, 1894.
2. *Les groupes réels simples et continus*, Annales de l'École Normale, vol. 31(1914), pp. 263–355.

M. Deuring.

1. *Algebren*, Berlin, 1935.

H. Hasse.

1. *Über p-adische Schiefkörper*, Math. Annalen, vol. 104(1931), pp. 495–534.
2. *Darstellbarkeit von Zahlen durch quadratische Formen in einem beliebigen algebraischen Zahlkörper*, Journal für Math., vol. 153(1924), pp. 113–130.
3. *Symmetrische Matrizen im Körper der rationalen Zahlen*, Journal für Math., vol. 153 (1924), pp. 12–43.
4. *Äquivalenz quadratischer Formen in einem beliebigen algebraischen Zahlkörper*, Journal für Math., vol. 153(1924), pp. 158–162.

N. Jacobson.

1. *A class of normal simple Lie algebras of characteristic zero*, Annals of Math., vol. 38(1937), pp. 508–517.
2. *Simple Lie algebras of type* A, Annals of Math., vol. 39(1938), pp. 181–188.
3. *A note on non-associative algebras*, this Journal, vol. 3(1937), pp. 544–548.
4. *A note on topological fields*, Am. Jour. of Math., vol. 59(1937), pp. 889–894.

W. Landherr.

1. *Über einfache Liesche Ringe*, Hamb. Abhandlungen, vol. 11(1935), pp. 41–64.
2. *Liesche Ringe vom Typus A*, Abhandlungen der Hansischen Univ., vol. 12(1938), pp. 200–241.

E. H. Moore.

1. *General Analysis*, I, Am. Philos. Soc. Publication, Philadelphia, 1935.

A. Weinstein.

1. *Fundamentalsatz der Tensorrechnung*, Math. Zeitsch., vol. 16(1923), pp. 78–91.

E. Witt.

1. *Theorie der quadratischen Formen in beliebigen Körpern*, Journal für Math., vol. 176(1936–1937), pp. 31–44.

University of North Carolina.

LINEAR FUNCTIONALS AND COMPLETELY ADDITIVE SET FUNCTIONS

By B. J. Pettis

Introduction. We should like first to recall a few well known definitions. Let $T = [t]$ be an abstract space composed of arbitrary elements t, and let $\mathcal{F}^T$ denote the collection of all subsets of T. If $\mathcal{F} = [F]$ is a non-vacuous subcollection of $\mathcal{F}^T$, then $\mathcal{F}$ is an *additive family*[1] if

(I) F in $\mathcal{F}$ implies $T - F$ in $\mathcal{F}$, and

(II) F_n in $\mathcal{F}$ for $n = 1, 2, \cdots$ implies $\sum_n F_n$ in $\mathcal{F}$.

A finite or denumerable aggregate δ of elements $F_1, \cdots, F_n, \cdots$ of $\mathcal{F}$ that are disjoint in pairs will be called a *split*; if δ has only a finite number of F_n's, it is a *finite split*. If Δ is the collection of all finite splits and Δ' that of all splits, then $\alpha(F)$ defined from $\mathcal{F}$ to the reals is *additive* if $\sum_\delta \alpha(F_i) = \alpha(\sum_\delta F_i)$ for every δ in Δ, and *completely additive* (c.a.) if this holds for every δ in Δ'. The *norm* of an additive $\alpha(F)$ is defined to be

$$(0.1) \qquad \| \alpha \| = \mathrm{Var}\,(\alpha, T) = \underset{\Delta}{\mathrm{l.u.b.}} \sum_\delta | \alpha(F_i) |.$$

For such an α it is clear that

$$(0.2) \qquad \underset{\mathcal{F}}{\mathrm{l.u.b.}} \,| \alpha(F) | \leqq \| \alpha \| \leqq 2(\underset{\mathcal{F}}{\mathrm{l.u.b.}}\, | \alpha(F) |),$$

and hence an additive α is bounded if and only if $\| \alpha \| < \infty$. If α is c.a., then $\| \alpha \| < \infty$ and α must be bounded.[2]

The space $A = [\alpha]$, consisting of the functions $\alpha(F)$ bounded and additive (b.a.) over $\mathcal{F}$, is, under the definition of norm given in (0.1), a linear normed complete space, i.e., a B-space;[3] this is likewise true of the subset $C = [\gamma]$ composed of the functions $\gamma(F)$ completely additive over $\mathcal{F}$. Thus these in particular are B-spaces: the collection A^T of functions b.a. over $\mathcal{F}^T$ and its subset C^T consisting of functions c.a. over $\mathcal{F}^T$.

The following notational rules will be generally obeyed, although not always:

(i) Roman capitals are subsets of T and script capitals are collections of such subsets.

Received March 14, 1938; presented to the American Mathematical Society, February 26, 1938.

[1] S. Saks, *Theory of the Integral*, Warsaw, 1937, p. 7. Hereafter we shall refer to this treatise as TI.

[2] TI, p. 10.

[3] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 53. The letters TOL will refer to this monograph.

(ii) Real-valued functions of *subsets* of T are denoted by small Greek letters, and collections of such functions by italicized capitals.

(iii) Small German letters represent real-valued functions of *elements* of T, while German capitals stand for aggregates composed of such functions.

(iv) Linear functionals (real-valued additive and continuous functions defined over a B-space) will be set in small italics.

(v) The symbol Λ is reserved for the empty set.

Under varying hypotheses on the space T, several authors[4] have given proofs of the following proposition:[5]

(M) *If* $\gamma(F)$ *is real-valued and c.a. over* $\mathfrak{F}^T$, *then* γ *vanishes identically if* $\gamma(\{t\}) = 0$ *for every set* $\{t\}$ *consisting of a single point t.*

If T has power $\leqq 2^{\aleph_0}$, (M) was proved by Banach and Kuratowski[6] using the hypothesis of the continuum. Later Ulam[7] showed that (M) *holds if*

(U) *the power of* T *is smaller than the first inaccessible*[8] *cardinal number*; and that from the assumption (which is weaker than the continuum hypothesis)

(J) *there is no inaccessible number* $\leqq 2^{\aleph_0}$,

it follows that if (M) is true for some T of power $\mathfrak{m}$, it is also true for all T of power $\leqq 2^{\mathfrak{m}}$. Thus if (J) is assumed, then (M) is, in particular, true for any T of power $2^{\aleph_0}$ and hence for the unit interval. In the present paper (M) is assumed in the majority of the theorems; in these the hypothesis (M) can, by Ulam's result, be replaced by (U).

In those spaces T for which (M) holds it is possible to give a slightly sharper form to the generalized Lebesgue decomposition theorem[9] concerning completely additive functions of "measurable" sets. This will be done in the present paper for both real- and abstract-valued functions, together with some applications to the linear functionals over certain spaces and to the relationships between c.a. functions and absolutely continuous[10] (a.c.) functions.

The first section is concerned with extending the range of b.a. functions and with the relative denseness of certain classes of c.a. functions with respect to corresponding classes of b.a. functions. In §2 some well known properties of l (the space of absolutely convergent series) are extended under the hypothesis

[4] S. Banach and C. Kuratowski, *Sur une généralisation du problème de la mesure*, Fund. Math., vol. 14(1929), pp. 127–131; S. Ulam, *Zur Masstheorie in der allgemeinen Mengenlehre*, Fund. Math., vol. 16(1930), pp. 140–150; E. Szpilrajn, *Remarques sur les fonctions complètement additives d'ensemble et sur les ensembles jouissant de la propriété de Baire*, Fund. Math., vol. 22(1934), pp. 303–311; W. Sierpinski and E. Szpilrajn, *Remarques sur le problème de la mesure*, Fund. Math., vol. 26(1936), pp. 256–261.

[5] The statement that the "generalized measure problem" (Banach and Kuratowski, loc. cit.) has a negative answer.

[6] Loc. cit.

[7] Loc. cit.

[8] A cardinal number $\aleph_\xi$ is *inaccessible* if ξ is a limit number, $\aleph_\xi > \aleph_0$, and $\aleph_\xi$ cannot be represented as a sum of less than $\aleph_\xi$ cardinal numbers each smaller than $\aleph_\xi$.

[9] TI, p. 33, Theorem 14.6.

[10] A function $\alpha(F)$, defined over a family of measurable sets measured by $\mu(F)$, is *absolutely continuous* (a.c.) if given $\epsilon > 0$ there is a $\delta > 0$ such that $|\alpha(F)| < \epsilon$ when $|\mu(F)| < \delta$.

(M) to the general space C^{T}, and in 2.3 the condition (M) is expressed in terms of the functionals linear over C^{T}. The decomposition theorem and some corollaries occupy §3, and in §4 some of the results of §1 and §3 are carried over to the case of abstract-valued functions.

1. Extension of the range of bounded additive functions; denseness theorems.

1.1. Theorem. *Let $\mathfrak{F}$ be an additive family and $\mu(\mathrm{F})$ a fixed function non-negative and c.a. over $\mathfrak{F}$. If $\mathfrak{F}' = [\mathrm{F}']$ is an additive subfamily of $\mathfrak{F}$ and $\alpha'(\mathrm{F}')$ is b.a. over $\mathfrak{F}'$ and satisfies the condition*

(N′) F′ *in* $\mathfrak{F}'$ *and* $\mu(\mathrm{F}') = 0$ *imply* $\alpha'(\mathrm{F}') = 0$,

then $\alpha'(\mathrm{F}')$ can be extended to form a function $\alpha(\mathrm{F})$ b.a. over $\mathfrak{F}$ and satisfying the conditions $\| \alpha \|_A = \| \alpha' \|_{A'}$ and

(N) F *in* $\mathfrak{F}$ *and* $\mu(\mathrm{F}) = 0$ *imply* $\alpha(\mathrm{F}) = 0$.

Let $\mathfrak{M}$ be the functions $\mathfrak{m}(t)$ that are μ-essentially bounded over T and are $\mathfrak{F}$-measurable;[11] when normed by $\| \mathfrak{m} \| = \mu\text{-ess. sup.} \mid \mathfrak{m}(t) \mid$, the space $\mathfrak{M}$ is a B-space. The subclass $\mathfrak{M}'$ of functions $\mathfrak{m}'(t)$ that are μ-essentially bounded over T and $\mathfrak{F}'$-measurable forms a closed linear subset of $\mathfrak{M}$. By a theorem of Hildebrandt and of Fichtenholz and Kantorovitch,[12] given the above α' the Radon-Stieltjes integral

$$f'(\mathfrak{m}') = \int_{\mathrm{T}} \mathfrak{m}'(t)\, d\alpha'$$

defines a linear functional f' over $\mathfrak{M}'$, and $\| f' \| = \| \alpha' \|_{A'}$. By the Hahn-Banach theorem on the extension of linear functionals, $f'(\mathfrak{m}')$ can be extended to form a linear functional $f(\mathfrak{m})$ over $\mathfrak{M}$, with $\| f \| = \| f' \|$. By the same Hildebrandt-Fichtenholz-Kantorovitch theorem there exists a function $\alpha(\mathrm{F})$ b.a. over $\mathfrak{F}$, satisfying the above condition (N), and defining f by means of the integral

$$f(\mathfrak{m}) = \int_{\mathrm{T}} \mathfrak{m}(t)\, d\alpha,$$

with $\| f \| = \| \alpha \|_A$. Hence $\| \alpha' \|_{A'} = \| \alpha \|_A$; and if we let $\mathfrak{m}_{\mathrm{F}'}$ be the characteristic function of an element F′ of $\mathfrak{F}'$, it follows that

$$\alpha(\mathrm{F}') = \int_{\mathrm{T}} \mathfrak{m}_{\mathrm{F}'}(t)\, d\alpha = f(\mathfrak{m}_{\mathrm{F}'}) = f'(\mathfrak{m}_{\mathrm{F}'}) = \int_{\mathrm{T}} \mathfrak{m}_{\mathrm{F}'}(t)\, d\alpha' = \alpha'(\mathrm{F}'),$$

so that $\alpha(\mathrm{F})$ is an extension of $\alpha'(\mathrm{F}')$. This completes the proof.

If we apply another result due to Hildebrandt[13] and to Fichtenholz and

[11] TI, p. 12.

[12] T. H. Hildebrandt, *On bounded linear functional operations*, Trans. Amer. Math. Soc., vol. 36(1934), pp. 868–875; especially p. 870. G. Fichtenholz and L. Kantorovitch, *Sur les opérations dans l'espace des fonctions bornées*, Studia Math., vol. 5(1934), pp. 69–98; especially p. 76.

[13] Loc. cit., p. 872.

Kantorovitch,[14] the following theorem, similar to the preceding, can be proved in the same fashion:

1.2. THEOREM. *If $\mathfrak{F}$ is an additive family and $\mathfrak{F}'$ is an additive subfamily of $\mathfrak{F}$, then any function α' that is b.a. over $\mathfrak{F}'$ can be extended to form a function $\alpha(\mathrm{F})$ b.a. over $\mathfrak{F}$ and having $\| \alpha \|_A = \| \alpha' \|_{A'}$.*

Using 1.1 we are able to state

1.3. THEOREM. *Let $\mathfrak{F}$ be an additive family and $\mu(\mathrm{F})$ a fixed function non-negative and c.a. over $\mathfrak{F}$. If there exists a denumerable split $\bar{\delta} = \{\mathrm{F}_1, \cdots, \mathrm{F}_n, \cdots\}$ with $\mu(\mathrm{F}_n) > 0$ for all n, then in the subspace A^{N} composed of elements α of A that satisfy condition (N) of 1.1 the set $C^{\mathrm{N}} = C \cdot A^{\mathrm{N}}$ is closed linear and nowhere dense, and hence is of the first category. This is also true of C with respect to A.*

It is easily verified that A^{N} and C^{N} are closed linear subsets of A; the proof will be omitted.

Let $\mathrm{S} = \sum_1^\infty \mathrm{F}_n$ and let $\mathfrak{F}'$ be the additive family consisting of the null set Λ adjoined to the collection composed of all sets obtained by adding a finite or denumerable number of the disjoint sets $\mathrm{F}_0 = \mathrm{T} - \mathrm{S}, \mathrm{F}_1, \cdots, \mathrm{F}_n, \cdots$.

If P is the set of positive integers, then in A^{P} the closed linear subset C^{P} is equivalent[15] to the separable B-space l consisting of real absolutely convergent series, so that C^{P} is a separable subset of A^{P}. But since A^{P} is the adjoint of the non-separable space formed by the bounded sequences of reals,[16] A^{P} must be non-separable.[17] Hence there exists a function π b.a. over $\mathfrak{F}^{\mathrm{P}}$ but not c.a.

For each F' in $\mathfrak{F}'$ define $\alpha'(\mathrm{F}') = \pi(\mathrm{R})$, where R is the element of $\mathfrak{F}^{\mathrm{P}}$ that consists of the positive integers n for which $\mathrm{F}_n \subset \mathrm{F}'$. Then α' is b.a. over $\mathfrak{F}'$ but not c.a.; and, moreover, α' satisfies condition (N′) since $\alpha'(\Lambda) = \alpha'(\mathrm{F}_0) = 0$, where Λ and F_0 are the only members of $\mathfrak{F}'$ that can possibly have μ-measure 0. From 1.1, α' can be extended over $\mathfrak{F}$ while preserving these properties. Hence $A^{\mathrm{N}} - C^{\mathrm{N}} \neq \Lambda$, which implies that $A - C \neq \Lambda$, so that the closed linear subspaces C^{N} and C must be nowhere dense in A^{N} and A, respectively.

1.31.[18] THEOREM. *If $\mathfrak{F}$ is an additive family containing a denumerable split $\bar{\delta} = \{\mathrm{F}_1, \cdots, \mathrm{F}_n, \cdots\}$ with $\mathrm{F}_n \neq \Lambda$ for each n, then C is closed linear and nowhere dense (and therefore of the first category) in A.*

Let $\sum_1^\infty \mu_n$ be a convergent series of positive numbers and for each n let t_n be a point in non-null F_n. The function $\mu(\mathrm{F}) = \sum_{t_n \in \mathrm{F}} \mu_n$ is non-negative and c.a.

[14] Loc. cit., p. 76.

[15] I.e., there exists a 1-1 additive and norm-preserving transformation mapping all of C^{P} onto all of l.

[16] Hildebrandt, loc. cit., p. 870; or see 2.2.

[17] TOL, p. 189, Theorem 12.

[18] Theorem 1.31 is not new since it follows from a much stronger result of Ulam (*Concerning functions of sets*, Fund. Math., vol. 14(1929), pp. 231–233; also see A. Tarski, *Une contribution à la théorie de la mesure*, Fund. Math., vol. 15(1930), pp. 42–50) stating that in $A^{\mathrm{P}} - C^{\mathrm{P}}$ there exists an element having 0 and 1 as its only functional values.

over $\mathfrak{F}$, and $\bar{\delta}$ is a split having $\mu(F_n) = \mu_n > 0$ for each n. The present theorem now follows immediately from 1.3.

2. **The space** C^T. In the remaining sections the additive family $\mathfrak{F}$ under discussion will always be understood to include the sets $\{t\}$ that consist of a single point; the elements of $\mathfrak{F}$ will sometimes be referred to as the *measurable sets*. If α is b.a. over such a family, then $|\alpha(\{t\})| > n^{-1}$ holds for at most a finite number of points t in T, so that $|\alpha(\{t\})| > 0$ is true for an at most denumerable set $S_\alpha = \{t_\alpha^i\}$ $(i = 1, 2, \cdots, n, \cdots)$ of points t; these we call the *spectral points of* α. The *spectral function* $\sigma_\alpha(F)$ *associate to* α is the element of A defined by

$$\sigma_\alpha(F) = \alpha(F \cdot S_\alpha),$$

and α is said to be a *spectral element of* A if $\alpha(F) \equiv \sigma_\alpha(F)$. If α is in A, then clearly $\alpha' = \alpha - \sigma_\alpha$ is also in A and $\alpha'(\{t\}) = 0$ for all points t; similarly, γ in C implies that $\gamma' = \gamma - \sigma_\gamma$ is in C and $\gamma'(\{t\}) = 0$ for all $\{t\}$. The set of spectral elements lying in C will be denoted by S; in particular, S^T stands for the set of spectral elements of A^T that are in C^T. It is evident that if γ is in S or S^T, then $\|\gamma\| = \|\sigma_\gamma\| = \sum_i |\gamma(\{t_\gamma^i\})|$. Proposition (M) can now be stated in the following form:

(M) *If* γ *is in* C^T, *then* $\gamma = \sigma_\gamma$;

or, even more briefly,

(M) $C^T = S^T$.

The next theorem is complementary to 1.2.

2.1. Theorem. *Suppose* (M) *holds. Then any* γ *in* C *can be extended to form an element* β *of* C^T *if and only if* γ *is in* S; *and the extension is unique. Conversely, each* β *in* C^T *is the extension of a unique* γ *in* S. *The norms of corresponding* γ *and* β *are equal.*

(M) implies that γ can be extended only if γ is spectral; on the other hand, if γ is spectral, then $\beta(F) = \gamma(F \cdot S_\gamma)$, $F \in \mathfrak{F}^T$, is an extension of γ. If β and β' are extensions of γ, then since $\mathfrak{F}$ contains $\{t\}$ for every point t, the function $\beta - \beta'$ vanishes for all $\{t\}$; from (M) it follows that $\beta - \beta' = 0$, and thus the extension is unique.

If β is in C^T, (M) implies that $\beta(F) = \beta(F \cdot S_\beta)$ for all F in $\mathfrak{F}^T$, so that $\gamma(F) = \beta(F \cdot S_\beta)$, $F \in \mathfrak{F}$, is in S and β is an extension of γ. Since a contraction is specified by its domain, γ is unique. Moreover, $\beta(\{t\}) = \gamma(\{t\})$ for all t, and since $\beta \in S^T$ and $\gamma \in S$, this is readily seen to imply the equality of the norms.

At this point we wish to quote a result due to Hildebrandt[19] concerning the linear functionals over the B-space $\mathfrak{B}^T$, the space composed of the real-valued functions $\mathfrak{b}(t)$ defined and bounded over T with $\|\mathfrak{b}\| = \operatorname{l.u.b.}_T |\mathfrak{b}(t)|$.

[19] The theorem is implicit in Hildebrandt's previously cited paper and explicitly stated in his *Linear operations on functions of bounded variation*, Bull. Amer. Math. Soc., vol. 44 (1938), p. 75. See also H. H. Goldstine, *Weakly complete Banach spaces*, this Journal, vol. 4(1938), pp. 125–131.

2.2. THEOREM (Hildebrandt). *A functional $f(\mathfrak{b})$ over $\mathfrak{B}^{\mathrm{T}}$ is linear if and only if there exists an α_f in A^{T} such that for all $\mathfrak{b}$ in $\mathfrak{B}^{\mathrm{T}}$*

$$f(\mathfrak{b}) = \int_{\mathrm{T}} \mathfrak{b}(t)\, d\alpha_f, \tag{2.21}$$

the integral being in the Radon-Stieltjes sense. The norm of f is[20] $\| \alpha_f \|$.

As a result complementary to 2.2 and as another generalization to C^{T} of a property of l, we have the following theorem, which also serves to state the generalized measure problem in terms of linear functionals.

2.3. THEOREM. *A necessary and sufficient condition that* (M) *hold for a given* T *is*:

(K) *A functional $f(\gamma)$ over C^{T} is linear if and only if there exists a $\mathfrak{b}_f$ in $\mathfrak{B}^{\mathrm{T}}$ such that*

$$f(\gamma) = \int_{\mathrm{T}} \mathfrak{b}_f(t)\, d\gamma \tag{2.31}$$

for all γ in C^{T}, and $\| f \| = \| \mathfrak{b}_f \|$.

Suppose an element $\mathfrak{b}$ of $\mathfrak{B}^{\mathrm{T}}$ is given. 2.2 implies that equation (2.31) for γ in A^{T} defines a linear functional $f^*(\gamma)$ over A^{T}. If f^* is considered only on C^{T}, it then defines a linear functional $f(\gamma)$ over C^{T}, and clearly

$$\| f \| \leqq \| f^* \| = \| \mathfrak{b} \|. \tag{2.32}$$

Now suppose $f(\gamma)$ is any functional linear over C^{T}; for each t in T define

$$\gamma_t(\mathrm{F}) = \begin{cases} 1 & \text{if } t \in \mathrm{F}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.33}$$

and let $\mathfrak{b}_f(t) = f(\gamma_t)$. Then $\mathfrak{b}_f$ is in $\mathfrak{B}^{\mathrm{T}}$, since

$$| \mathfrak{b}_f(t) | \leqq \| f \| \cdot \| \gamma_t \| = \| f \|, \tag{2.34}$$

and hence $\mathfrak{b}_f$ defines a linear functional

$$g(\gamma) = \int_{\mathrm{T}} \mathfrak{b}_f(t)\, d\gamma$$

over C^{T}, with the equality

$$g(\gamma_t) = \mathfrak{b}_f(t) = f(\gamma_t) \tag{2.35}$$

holding for all t in T, and

$$\| g \| \leqq \| \mathfrak{b}_f \| \leqq \| f \| \tag{2.36}$$

by (2.32) and (2.34).

[20] A functional $f(\mathfrak{b})$ linear over B^{T} is *positive* if $\mathfrak{b}(t) \geqq 0$ for all t implies $f(\mathfrak{b}) \geqq 0$; cf. R. P. Bailey, *Convergence of sequences of positive linear functional operations*, this Journal, vol. 2 (1936), pp. 287–303. These functionals are clearly those defined by the non-negative elements of A^{T}, i.e., by the functions $\alpha(\mathrm{F})$ bounded and additive over $\mathfrak{F}^{\mathrm{T}}$ and such that $\alpha(\mathrm{F}) \geqq 0$ for all F.

But if (M) holds for T, then the set of finite linear combinations of the elements γ_t is dense in C^{T}; since the linear functionals g and f coincide over this set by (2.35), we must have $g = f$ and

$$f(\gamma) = \int_{\mathrm{T}} \mathfrak{b}_f(t)\, d\gamma,$$

where $\mathfrak{b}_f(t) = f(\gamma_t)$; and from (2.36)

$$\| f \| = \| g \| = \| \mathfrak{b}_f \| .$$

This completes the proof of the necessity of (K).

Conversely, suppose (K) holds, i.e., $\mathfrak{B}^{\mathrm{T}}$ is adjoint to C^{T}. The linear functionals

$$f_t(\mathfrak{b}) = \mathfrak{b}(t)$$

defined over $\mathfrak{B}^{\mathrm{T}}$ form a total set of functionals, that is, $f_t(\mathfrak{b}) = 0$ for all t implies $\mathfrak{b} = 0$; and these functionals are evidently those defined by the elements γ_t of C^{T} given in (2.33). Thus if $\mathfrak{B}^{\mathrm{T}}$ is adjoint to C^{T}, then the γ_t's form in C^{T} a total set of elements—if $b(\gamma)$ is linear over C^{T} and $b(\gamma_t) = 0$ for all t, then $b(\gamma) \equiv 0$. The set of γ_t's must then[21] have the set of its finite linear combinations dense in C^{T}. But the limit of a sequence of such combinations will be zero on any set which is disjoint with the denumerable set of points consisting of all the spectral points of the members of the sequence. Thus (M) holds.

2.4. Corollary. *If* T *has its power smaller than the first inaccessible number, then $f(\gamma)$ is a linear functional over C^{T} if and only if there is an element $\mathfrak{b}_f$ in $\mathfrak{B}^{\mathrm{T}}$ such that*

$$f(\gamma) = \int_{\mathrm{T}} \mathfrak{b}_f(t)\, d\gamma$$

and $\| f \| = \| \mathfrak{b}_f \|$ *; in particular this holds for* $\mathrm{T} = [0, 1]$, *under the hypothesis* (J).

The next two theorems extend to C^{T} two more well known properties of l.

2.5. Theorem. *Suppose* (M) *holds. If γ_n is in C^{T} $(n = 1, 2, \cdots)$ and $\lim_{n\to\infty} \gamma_n(\mathrm{F}) = \gamma_0(\mathrm{F})$ exists for every* F *in $\mathfrak{F}^{\mathrm{T}}$, then γ_0 is in C^{T}, $\lim_n \gamma_n = \gamma_0$ in C^{T}, and $\lim_n \gamma_n(\mathrm{F}) = \gamma_0(\mathrm{F})$ exists uniformly in* F. *Hence under* (M) *weak convergence in C^{T} implies strong, or norm, convergence.*

That γ_0 is in C^{T} is due to a result of Nikodym's.[22] Since (M) holds, we have $\gamma_n = \sigma_{\gamma_n}$ $(n = 0, 1, \cdots, j, \cdots)$; hence it is only necessary to show that $\lim_n \sigma_{\gamma_n} = \sigma_{\gamma_0}$. Let G be the denumerable set consisting of all the spectral points of the γ_n's $(n = 0, 1, \cdots, j, \cdots)$. Since for every F

$$\lim_{n\to\infty} [\sigma_{\gamma_n}(\mathrm{F}) - \sigma_{\gamma_0}(\mathrm{F})] = 0,$$

[21] TOL, p. 58, Theorem 7.

[22] *Sur les suites des fonctions parfaitment additives d'ensembles abstraits*, Comptes Rendus, vol. 192(1931), pp. 727–728.

and each σ_{γ_n} is c.a., it follows that

$$\lim_{n\to\infty} [\sum_{t \in G'} \gamma_n(\{t\}) - \sum_{t \in G'} \gamma_0(\{t\})] = 0$$

for every subset G′ of G. This, by a result of Schur's,[23] implies

$$\lim_{n\to\infty} \sum_{t \in G} | \gamma_n(\{t\}) - \gamma_0(\{t\}) | = 0,$$

and hence $\lim_{n\to\infty} || \sigma_{\gamma_n} - \sigma_{\gamma_0} || = 0$, or $\lim_n \gamma_n = \gamma_0$ in C^T. It is a consequence of this and of (0.2) that $\lim_n \gamma_n(F) = \gamma_0(F)$ uniformly in F.

If $\{\gamma_n\}$ is a weakly convergent sequence in C^T and $\mathfrak{m}_F(t)$ is the characteristic function of an arbitrary F, then by 2.3

$$\lim_{n\to\infty} \gamma_n(F) = \lim_{n\to\infty} \int_T \mathfrak{m}_F(t)\, d\gamma_n = \int_T \mathfrak{m}_F(t)\, d\gamma_0 = \gamma_0(F)$$

exists for every F in $\mathcal{F}^T$. By the first part of the theorem this implies that $\lim_n \gamma_n = \gamma_0$ strongly in C^T.

2.6. THEOREM. *If* (M) *holds, then every function* X(*s*) *of bounded variation from a linear interval* I = $[s_1, s_2]$ *to* C^T *is strongly differentiable a.e. in* I.

By a theorem of Dunford and Morse[24] it is sufficient to prove that any function X(*s*), defined from I to C^T and satisfying a Lipschitz condition, is strongly differentiable almost everywhere in I. Let $\{r_i\}$ be the rationals in I, and let X be the separable closed linear manifold in C^T generated by the elements $\{X(r_i)\}$; since X(*s*) is continuous, the range of X(*s*) lies in X. But X is obtained by taking finite linear combinations of the $X(r_i)$ and closing. Since neither of these operations adds spectral points, it follows that the set consisting of the spectral points of all the elements of X is at most denumerable. Hence the spectral elements associated to the elements of X can be put in a 1-1 additive and norm-preserving correspondence with a subset Y of the space l. Since (M) implies that each element of X is its own spectral function, the closed linear manifold X can be mapped onto Y by a 1-1 additive and norm-preserving correspondence K(*x*). The function Y(*s*) = K(X(*s*)) from I to Y evidently exists, since X(*s*) has its range in X. It now follows from the properties of K(*x*) that (1) Y(*s*) satisfies a Lipschitz condition and is therefore, by a theorem due to Clarkson,[25] strongly differentiable a.e. in I, and hence (2) X(*s*) is strongly differentiable a.e. in I. This ends the proof.

3. The decomposition of completely additive functions.

In this section the fixed family $\mathcal{F}$ is supposed to be complete[26] with respect to a fixed non-negative

[23] TOL, p. 138.

[24] *Remarks on the preceding paper of James A. Clarkson*, Trans. Amer. Math. Soc., vol. 40(1936), pp. 415–420; especially p. 415.

[25] *Uniformly convex spaces*, Trans. Amer. Math. Soc., vol. 40(1936), pp. 396–414; especially p. 412.

[26] TI, p. 86.

element $\mu(\mathrm{F})$ of C that vanishes for single points. Under these circumstances an extended Lebesgue integral can be defined[27] in terms of $\mathcal{F}$ and μ; the functions $\mathfrak{c}(t)$ integrable by this definition will be denoted by $\mathfrak{L}_\mu$ and their integrals (which form a subset of C) by L. The symbol C^0 will represent the elements γ of C that have the property $\gamma(\{t\}) = 0$ for all t.

3.1. Theorem. *Suppose* (M) *holds. Then if γ is in C, there is a unique β_γ in C^{T} and a unique $\mathfrak{c}_\gamma$ in $\mathfrak{L}_\mu$ such that*

$$\gamma(\mathrm{F}) = \beta_\gamma(\mathrm{F}) + \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\, d\mu \tag{3.11}$$

for every measurable F. *Here β_γ is the extension of the spectral function σ_γ of γ, and hence*

$$\begin{aligned} \gamma(\mathrm{F}) &= \gamma(\mathrm{F}\cdot \mathrm{S}_\gamma) + \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\, d\mu \\ &= \sum_{t_\gamma^i \in \mathrm{F}} \gamma(\{t_\gamma^i\}) + \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\, d\mu, \end{aligned} \tag{3.12}$$

where $\mathrm{S}_\gamma = \{t_\gamma^i\}$ *is the at most denumerable set of points t for which* $\gamma(\{t\}) \neq 0$. *Moreover,*

$$\|\gamma\| = \sum_i |\gamma(\{t_\gamma^i\})| + \int_{\mathrm{T}} |\mathfrak{c}_\gamma(t)|\, d\mu = \|\sigma_\gamma\| + \|\mu_\gamma\| \tag{3.13}$$

where $\mu_\gamma(\mathrm{F}) = \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\, d\mu$.

The family $\mathcal{F}$ being additive and complete with respect to μ, for each γ in C there is a set H_γ of μ-measure zero and a $\mathfrak{c}_\gamma$ in $\mathfrak{L}_\mu$ such that[28]

$$\begin{aligned} \gamma(\mathrm{F}) &= \gamma(\mathrm{F}\cdot \mathrm{H}_\gamma) + \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\, d\mu \\ &= \gamma(\mathrm{F}\cdot \mathrm{H}_\gamma) + \mu_\gamma(\mathrm{F}). \end{aligned}$$

The function $\mathfrak{c}_\gamma(t)$ may be supposed to have all its values finite. Since $\mathcal{F}$ is complete and H_γ has measure zero, the function

$$\beta_\gamma(\mathrm{E}) \equiv \gamma(\mathrm{E}\cdot \mathrm{H}_\gamma)$$

is defined and c.a. over $\mathcal{F}^{\mathrm{T}}$; hence there exists a decomposition of the form (3.11). The measure function μ being in C^0, it follows that μ_γ is also, since $\mu_\gamma(\{t\}) = \mathfrak{c}_\gamma(t)\cdot\mu(\{t\})$; in any decomposition (3.11) it is then true that $\gamma(\{t\}) =$

[27] O. Nikodym, *Sur une généralisation des intégrales de M. J. Radon*, Fund. Math., vol. 15(1930), pp. 130–179; or Chapter I of TI.

[28] TI, p. 33, Theorem 14.6.

$\beta_\gamma(\{t\})$, implying from 2.1 that β_γ in C^T must be the unique extension of σ_γ. Hence

$$\begin{aligned}\gamma(F) &= \sigma_\gamma(F) + \mu_\gamma(F) \\ &= \gamma(F\cdot S_\gamma) + \int \mathfrak{c}_\gamma(t)\, d\mu \\ &= \sum_{t_\gamma^i \in F} \gamma(\{t_\gamma^i\}) + \int_F \mathfrak{c}_\gamma(t)\, d\mu,\end{aligned}$$

and this establishes (3.12). The uniqueness of $\mathfrak{c}_\gamma$ follows from that of β_γ.

The equality between norms results from

$$\begin{aligned}\| \gamma \| &= \mathrm{Var}(\gamma, S_\gamma) + \mathrm{Var}(\gamma, T - S_\gamma) \\ &= \| \sigma_\gamma \| + \| \mu_\gamma \|.\end{aligned}$$

3.2. Corollary. *Under the hypothesis* (J) *any* $\gamma(F)$ *c.a. over the Lebesgue-measurable sets in* [0, 1] *can be decomposed into*

$$\gamma(F) = \sum_{t^i \in F} \gamma(\{t^i\}) + \int_F \mathfrak{c}_\gamma(t)\, dt,$$

where $S_\gamma = \{t^i\}$ *is the at most denumerable set of points* t *for which* $\gamma(\{t\}) \neq 0$.

3.3. Theorem. *Under* (M) *the subspaces* C^0, C^N, *and* L *are all identical.*[29]

That $C^N = L$ has been proved by Nikodym.[30] The identity between C^0 and L follows from (3.12).

3.31. Corollary. *Under* (M), *if* $\mathfrak{F}$ *is an additive family completed with respect to an element* μ *of* C^0 *and* $\mathfrak{F}$ *is separable when metrized by*

$$\text{(3.311)} \qquad \mathrm{dist}\,(F, G) = \mu(F - G) + \mu(G - F),$$

then $\mathfrak{F}$ *is separable when so metrized by any other element* ν *of* C^0.

If ν is in C^0, then by 3.3 it is a μ-integral and hence is a.c. with respect to μ.

3.32. Corollary. *Under* (J) *a c.a. function of Lebesgue-measurable sets in* [0, 1] *is a.c. if it vanishes for single points; hence it is a.c. if it is continuous.*[31]

3.4. Theorem. *Under* (M) *a functional* $f(\gamma)$ *over* C *is linear if and only if there is a* $\mathfrak{b}_f$ *in* $\mathfrak{B}^T$ *and an* $\mathfrak{m}_f$ *in* $\mathfrak{M}$ *such that for all* γ *in* C

$$\begin{aligned}f(\gamma) &= \int_T \mathfrak{b}_f(t)\, d\beta_\gamma + \int_T \mathfrak{m}_f(t)\mathfrak{c}_\gamma(t)\, d\mu \\ &= \sum_{S_\gamma} \mathfrak{b}_f(t)\cdot\gamma(\{t\}) + \int_T \mathfrak{m}_f(t)\mathfrak{c}_\gamma(t)\, d\mu,\end{aligned}$$

[29] This seems to imply that when $\mathfrak{F}$ is completed with respect to a function μ, non-negative and c.a. over $\mathfrak{F}$ and vanishing for single points, then all measure functions c.a. over the *new, completed* family and vanishing for single points have precisely the same sets of measure zero.

[30] O. Nikodym, loc. cit., in footnote 27, Theorem IV, p. 179; or TI, p. 36, Theorem 14.11.

[31] A function $\eta(F)$ defined over the L-measurable sets in [0, 1] is *continuous* if $\lim \eta(F) = 0$ as $\delta(F)$, the diameter of F, tends to 0.

where $\gamma(\mathrm{F}) = \beta_\gamma(\mathrm{F}) + \int_{\mathrm{F}} \mathfrak{c}_\gamma(t)\,d\mu$, β_γ *is in* C^{T}, *and* $\mathfrak{c}_\gamma$ *is in* $\mathfrak{L}_\mu$. *The elements* $\mathfrak{b}_f$ *and* $\mathfrak{m}_f$ *are unique and*

$$\|f\| = \max[\|\mathfrak{b}_f\|, \|\mathfrak{m}_f\|].$$

If we write γ in C uniquely as

$$\gamma = \sigma_\gamma + \mu_\gamma \tag{3.30}$$

with σ_γ in S and μ_γ in L, then

$$\begin{aligned} f(\gamma) &= f(\sigma_\gamma) + f(\mu_\gamma) \\ &= f_1(\sigma_\gamma) + f_2(\mu_\gamma), \end{aligned}$$

where f_1 and f_2 are linear over the B-spaces S and L, respectively, and

$$\|f\| = \underset{\|\gamma\|=1}{\mathrm{l.u.b.}} |f(\gamma)| \geqq \|f_1\|, \|f_2\|.$$

On the other hand,

$$\begin{aligned} |f(\gamma)| \leqq |f_1(\sigma_\gamma)| + |f_2(\mu_\gamma)| &\leqq \|f_1\|\cdot\|\sigma_\gamma\| + \|f_2\|\cdot\|\mu_\gamma\| \\ &\leqq [\|\sigma_\gamma\| + \|\mu_\gamma\|]\cdot\max_{i=1,2}\|f_i\| = \|\gamma\|\cdot\max_{i=1,2}\|f_i\|, \end{aligned}$$

by (3.13). Hence $\|f\| = \max_{i=1,2}\|f_i\|$.

From 2.1 and 2.3 there exists a $\mathfrak{b}_f$ in $\mathfrak{B}^{\mathrm{T}}$ such that $\|f_1\| = \|\mathfrak{b}_f\|$ and

$$f_1(\sigma_\gamma) = \int_{\mathrm{T}} \mathfrak{b}_f(t)\,d\beta_\gamma,$$

where β_γ in C^{T} is the unique extension of σ_γ; hence

$$f_1(\sigma_\gamma) = \sum_{\mathrm{S}_\beta} \mathfrak{b}_f(t)\beta_\gamma(\{t\}) = \sum_{\mathrm{S}_\beta} \mathfrak{b}_f(t)\sigma_\gamma(\{t\}) = \sum_{\mathrm{S}_\gamma} \mathfrak{b}_f(t)\sigma_\gamma(\{t\}).$$

Similarly, $f_2(\mu_\gamma)$ being linear over L, there is a measurable and μ-essentially bounded $\mathfrak{m}_f(t)$ such that[32]

$$f_2(\mu_\gamma) = \int_{\mathrm{T}} \mathfrak{m}_f(t)\mathfrak{c}_\gamma(t)\,d\mu,$$

with $\|f_2\| = \|\mathfrak{m}_f\|$.

The uniqueness of $\mathfrak{m}_f$ and $\mathfrak{b}_f$ follows from that of the decomposition (3.30).

4. **Abstract-valued functions.** We now consider the extensions of the theorems of §1 and §3 to the case of functions $x(\mathrm{F})$ defined over $\mathcal{F}$ and taking their values in a fixed B-space X.

The definitions of *additive* and *c.a.* functions given in the introduction are formally retained. However, the definition of the norm of an additive function

[32] O. Nikodym, *Contribution à la théorie des fonctionelles linéaires*, etc., Mathematica, vol. 5(1931), pp. 130–141, Theorem I.

must be altered if the closest parallelism to the real-valued case is to be reached. Let $\bar{X} = [f]$ be the space of linear functionals over X; then the *norm*, $\|x\|$, *of the function* $x(F)$ *additive from* $\mathcal{F}$ *to* X is defined to be

$$(4.0) \qquad \|x\| = \underset{\|f\|=1}{\text{l.u.b.}} \|f(x(\cdot))\|_A,$$

where $f(x(\cdot))$ represents the function $f(x(F))$ considered as an element of the space of additive real-valued functions. From (4.0) and (0.2) it follows that for an additive $x(F)$

$$\|x(F)\|_X = \underset{\|f\|=1}{\text{l.u.b.}} |f(x(F))| \leqq \underset{\|f\|=1}{\text{l.u.b.}} \|f(x(\cdot))\|_A = \|x\|,$$

and on the other hand

$$\|x\| \leqq \underset{\|f\|=1}{\text{l.u.b.}} (2\cdot\underset{\mathcal{F}}{\text{l.u.b.}} |f(x(F))|) \leqq \underset{\|f\|=1}{\text{l.u.b.}} (2\cdot\underset{\mathcal{F}}{\text{l.u.b.}} \|x(F)\|_X),$$

whence

$$(4.1) \qquad \underset{\mathcal{F}}{\text{l.u.b.}} \|x(F)\|_X \leqq \|x\| \leqq 2\cdot\underset{\mathcal{F}}{\text{l.u.b.}} \|x(F)\|_X.$$

An immediate corollary is that *additive* $x(F)$ *is bounded if and only if* $\|x\| < \infty$.

As in the real-valued case, *a c.a. function* $x(F)$ *must be bounded.* Suppose $x(F)$ is c.a. Then

$$U(f) = f(x(\cdot))$$

is an additive operation from $\bar{X}$ to C; moreover, since the operation U is closed, i.e.,

$$\lim_{n\to\infty} \|f_n - f_0\| = 0 \quad \text{and} \quad \lim_{n\to\infty} \|U(f_n) - \gamma_0\| = 0 \quad \text{imply} \quad \gamma_0 = U(f_0),$$

U must be linear.[33] Then $\underset{\|f\|=1}{\text{l.u.b.}} \|f(x(\cdot))\|_C = \|U\| < \infty$, and hence $x(F)$ is bounded.[34]

Thus the linear space A_X, consisting of the functions $x(F)$ additive and bounded from $\mathcal{F}$ to X, is also normed, by (4.0), and the c.a. functions form a normed linear subspace C_X. These two spaces are also complete, i.e., A_X *and* C_X *are* B-*spaces.*[35] If $\lim_{m,n\to\infty} \|x_m - x_n\| = 0$, then from the completeness of X and from (4.1),

$$\lim_{n\to\infty} x_n(F) = x_0(F)$$

[33] TOL, p. 41, Theorem 7.

[34] What has actually been proved is that if $x(F)$ is additive over $\mathcal{F}$ and $f(x(F))$ is bounded for every f in $\bar{X}$, then $x(F)$ is bounded. This is a particular case of a more general theorem the proof of which we have borrowed in the above; see N. Dunford, *Uniformity in linear spaces*, to appear in vol. 44 of the Trans. Amer. Math. Soc.

[35] This has been shown for C_X using the norm $\underset{\mathcal{F}}{\text{l.u.b.}} \|x(F)\|$; cf. G. Birkhoff, *Integration of functions with values in a Banach space*, Trans. Amer. Math. Soc., vol. 38(1935), pp. 357–378; Theorem 10.

exists uniformly in F. The limit function $x_0(\mathrm{F})$ is clearly additive, and from the uniformity in F and the boundedness of each x_n it also must be bounded; moreover, since $x_0(\mathrm{F})$ is the uniform limit of $\{x_n(\mathrm{F})\}$,

$$\lim_{n\to\infty} \| x_n - x_0 \| = 0,$$

by (4.1). If the x_n's are in C_{X}, then from the uniform convergence over $\mathfrak{F}$ the limit function x_0 must be c.a. also. Thus A_{X} and C_{X} are both complete.

The subspaces $A_{\mathrm{X}}^{\mathrm{N}}$ and $C_{\mathrm{X}}^{\mathrm{N}}$ corresponding to the A^{N} and C^{N} of Theorem 1.3 are also B-spaces, and Theorem 1.3 still holds. It is sufficient to take some real-valued $\alpha_0(\mathrm{F})$ in $A^{\mathrm{N}} - C^{\mathrm{N}}$ and an x_0 in X with $\| x_0 \| \neq 0$; then $x_0\cdot\alpha_0(\mathrm{F})$ is in $A_{\mathrm{X}}^{\mathrm{N}} - C_{\mathrm{X}}^{\mathrm{N}}$, and closed linear $C_{\mathrm{X}}^{\mathrm{N}}$ is nowhere dense in $A_{\mathrm{X}}^{\mathrm{N}}$. The same method is applicable to Theorem 1.31.

If $x(\mathrm{F})$ *is in* C_{X}, *then,* as with real-valued functions, $x(\{t\}) \neq 0$ *for an at most denumerable set* S_x *of points* t; hence *a c.a.* $x(\mathrm{F})$ *has a spectral function* $x_{\mathrm{S}}(\mathrm{F})$. This, however, is not always true for the b.a. functions: the function $x(\mathrm{F})$ from $\mathfrak{F}^{\mathrm{T}}$ to $\mathfrak{B}^{\mathrm{T}}$, where T is the unit interval and the point $x(\mathrm{F})$ in $\mathfrak{B}^{\mathrm{T}}$ is the characteristic function of the set F, is bounded and additive; yet $\| x(\{t\}) \| = 1$ for every t.

It is unnecessary to redefine (M) in terms of $C_{\mathrm{X}}^{\mathrm{T}}$; the new (M) is implied by the old. Suppose x is in $C_{\mathrm{X}}^{\mathrm{T}}$ but $x - x_{\mathrm{S}} \neq 0$. Then there is an F′ in $\mathfrak{F}^{\mathrm{T}}$ for which $\| x(\mathrm{F}') - x_{\mathrm{S}}(\mathrm{F}') \| \neq 0$, and hence an f' exists in $\bar{\mathrm{X}}$ such that

$$| f'(x(\mathrm{F}') - x_{\mathrm{S}}(\mathrm{F}')) | = \| x(\mathrm{F}') - x_{\mathrm{S}}(\mathrm{F}') \| \neq 0.$$

But $f'(x - x_{\mathrm{S}})$ vanishes for single points and $f'(x - x_{\mathrm{S}})$ is c.a., and yet $f'(x - x_{\mathrm{S}}) \neq 0$; this contradicts (M).

Thus 2.1 carries over. In 3.1 two alterations are made; $\mathfrak{F}$ is supposed separable with respect to μ, and the μ-integral term in (3.11) and (3.12) becomes a set function that is a.c. with respect to μ. The proof is short. The function $y(\mathrm{F}) = x(\mathrm{F}) - x_{\mathrm{S}}(\mathrm{F})$ is c.a. and vanishes for single points, i.e., is in C_{X}^{0}; hence for each f in $\bar{\mathrm{X}}$ the function $f(y(\mathrm{F}))$ is in C^0, and by 3.3 must therefore be a.c. with respect to μ. Since $\mathfrak{F}$ is separable, by a theorem of Dunford's[36] it follows that $y(\mathrm{F})$ is a.c. with respect to μ. We then have the theorem: *under* (M), *if* $\mathfrak{F}$ *is completed with respect to* μ *and* $\mathfrak{F}$ *is separable when metrized according to* (3.311), *then any function c.a. from* $\mathfrak{F}$ *to a* B*-space* X *can be decomposed into*

$$x(\mathrm{F}) = x_{\mathrm{S}}(\mathrm{F}) + y(\mathrm{F}) = \sum_{t_x^i \in \mathrm{F}} x(\{t_x^i\}) + y(\mathrm{F}), \tag{4.2}$$

where $x_{\mathrm{S}}(\mathrm{F})$ *is the spectral function associated to* $x(\mathrm{F})$, $\mathrm{S}_x = \{t_x^i\}$ *is the at most denumerable set of points* t *for which* $\| x(\{t\}) \| \neq 0$, *and* $y(\mathrm{F})$ *is a.c. with respect to* μ *and c.a.* The $y(\mathrm{F})$ term in this decomposition cannot be improved, at least in the direction of certain integrals: there exists a function c.a. and a.c.

[36] Loc. cit., footnote 35.

from the L-measurable sets in [0, 1] to Hilbert space that fails to be an integral under any of several extensions of the Lebesgue.[37]

Thus 3.2 and 3.3 also hold in the abstract case, provided that μ-integrals are replaced by functions that are c.a. and a.c. with respect to μ. Corollary 3.31 has its evident analogue, and 3.32 becomes: *under* (J) *a c.a. function from the L-measurable sets in* [0, 1] *to a* B-*space* X *is a.c. if it vanishes for single points; hence it is a.c. if it is continuous.*

UNIVERSITY OF VIRGINIA.

[37] See *On integration in vector spaces*, to appear in vol. 44 of the Trans. Amer. Math. Soc.

SOLUTIONS OF SYSTEMS OF DIFFERENTIAL EQUATIONS IN INFINITELY MANY UNKNOWNS BY INFINITE SERIES OF DEFINITE INTEGRALS

BY JESSE PIERCE

Introduction. The method used in two papers by the author[1,2] can be extended to systems of differential equations in infinitely many dependent variables, but the integrands of the definite integrals appearing in the solutions consist of an infinite number of terms. F. R. Moulton has also examined the case where the right members of the differential equations

$$\frac{dx_i}{dt} = f_i(t, x_j) \qquad (i, j = 1, 2, \cdots) \tag{1}$$

are analytic in the variables t and x_j and $f_i(0, 0, \cdots) = 0$.[3]

In all three of these methods the solutions are given in terms of infinite series each term of which is an infinite series. The methods are theoretically correct, but as each term depends upon the preceding terms, it is clear that a serious practical obstacle is encountered when the terms after the first ones are to be computed.

Other known methods of approach to the problem are general analysis and infinite matrices.

In the present paper the solutions are given in terms of infinite series of definite integrals each integrand of which has a finite number of terms.

The system of differential equations to be considered has the form

$$\frac{dx_i}{dt} = \theta_i(t) + \sum_{h=1}^{\infty} f_{i\mu_1\cdots\mu_\nu}(t) x_1^{\mu_1} x_2^{\mu_2} \cdots x_\nu^{\mu_\nu} \qquad (i, \nu = 1, 2, \cdots), \tag{2}$$

where $\mu_1, \cdots, \mu_\nu$ are non-negative integers and

$$h = \mu_1 + 2\mu_2 + \cdots + \nu\mu_\nu, \tag{3}$$

for all positive integral values of ν. The above arrangement of the terms in the right members of (2) is made in order that every term with a finite number of the x_ν with finite exponents will appear after a finite number of terms. The path of integration for the definite integrals appearing in the solution functions is the interval (t_0, t), where the variable t is real and

$$|t_0 - t| = \mu. \tag{4}$$

Received March 26, 1938.

[1] This Journal, vol. 3(1937), pp. 616–622.

[2] American Mathematical Monthly, vol. 43(1936), pp. 530–539.

[3] F. R. Moulton, *Differential Equations*, p. 375.

For convenience the sequence $\mu_1, \cdots, \mu_\nu$ will be represented by μ, and hence the coefficient functions $f_{i\mu_1,\cdots,\mu_\nu}(t)$ will be represented by $f_{i\mu}(t)$. The functions $\theta_i(t)$, $f_{i\mu}(t)$ are real functions of the real variable t which are integrable (Riemann), but not necessarily continuous or bounded, on the interval (t_0, t) and satisfy the following conditions.

I. There exists a real positive function $\theta(u)$ such that

$$| \theta_i(t) | \leqq \theta(u), \qquad | f_{i\mu}(t) | \leqq A_\mu \theta(u) \qquad (i = 1, 2, \cdots), \tag{5}$$

where the A_μ are real positive constants and the function $\theta(u)$ is integrable (Riemann) on the interval (t_0, t).

II. The series

$$1 + \sum_{h=1}^{\infty} A_\mu X^{\mu_1 + \cdots + \mu_\nu} \qquad (\nu = 1, 2, \cdots) \tag{6}$$

converges and has the upper bound M when

$$| X | \leqq R. \tag{7}$$

III. The function $\theta(u)$ satisfies the inequality

$$\int_0^u \theta(u) + c = U(u) \leqq \frac{R}{2M}, \tag{8}$$

where c is a non-negative constant, and

$$\frac{d}{du} \int_0^u \theta(u)\, du = \theta(u), \tag{9}$$

except possibly at the set of points E of discontinuity of the functions $\theta_i(t)$, $f_{i\mu}(t)$.

Formal solutions of the system of differential equations (2) are found in §1, and these series are proved to converge in §2. In §3 the formal solutions are proved to have derivatives at all points on the interval (t_0, t), except at the set of points E.

1. **Formal solution of the system of differential equations** (2). The transformation

$$x_i = \sum_{m=1}^{\infty} K^{m+i-1} y_{im} \qquad (i = 1, 2, \cdots), \tag{10}$$

where K is an arbitrary parameter, reduces equations (2) to the form

$$\begin{aligned} \sum_{m=1}^{\infty} K^{m+i-1} \frac{dy_{im}}{dt} &= \theta_i(t) + f_{i1}(t) y_{11} K \\ &\quad + \{ f_{i1}(t) y_{12} + f_{i01}(t) y_{21} + f_{i2}(t) y_{11}^2 \} K^2 + \cdots \\ &= \theta_i(t) + \sum_{m=2}^{\infty} [\sum \alpha_{m\mu'} f_{i\mu}(t)[11] \cdots [1m_1] \cdots [\nu 1] \cdots [\nu m_\nu]] K^{m-1}, \end{aligned} \tag{11}$$

where $[pq]$ denotes y with subscripts pq and exponent μ_{pq}; μ' represents the sequence $\mu_{11}, \cdots, \mu_{1m_1}, \cdots, \mu_{\nu 1}, \cdots, \mu_{\nu m_\nu}$; the $m_1, \cdots, m_\nu$ are positive integers less than m; the $\alpha_{m\mu'}$ are positive constants; and $\sum$ represents the sum of all of the terms for which

$$(12)\quad \mu_{11} + 2\mu_{12} + \cdots + m_1\mu_{1m_1} + \cdots + \nu\mu_{\nu 1} + \cdots + (\nu + m_\nu - 1)\mu_{\nu m_\nu} = m - 1.$$

The parameter K is introduced in order that a convenient arrangement of the terms of the right members of (11) can be made.

A formal solution of the system of differential equations (11) can be found by replacing K by unity and solving, sequentially, the system

$$(13)\quad \begin{cases} \dfrac{dy_{i1}}{dt} = \theta_i(t), \\ \dfrac{dy_{i2}}{dt} = f_{i1}(t)y_{11}, \\ \dfrac{dy_{i3}}{dt} = f_{i1}(t)y_{12} + f_{i01}(t)y_{21} + f_{i2}(t)y_{11}^2, \\ \dfrac{dy_{im}}{dt} = \sum \alpha_{m\mu'} f_{i\mu}(t)[11] \cdots [1m_1] \cdots [\nu 1] \cdots [\nu m_\nu] \qquad (m = 2, 3, \cdots). \end{cases}$$

Equations (13) are obtained by equating the coefficient of K^{m+i-1} in the left member of (11) to the coefficient of K^{m-1} in the right member. The system of differential equations (13) has the solution

$$(14)\quad \begin{cases} y_{i1} = \displaystyle\int_{t_0}^{t} \theta_i(t)\,dt + c_i \equiv \eta_{i1}(t), \\ y_{i2} = \displaystyle\int_{t_0}^{t} f_{i1}(t)\eta_{11}(t)\,dt \equiv \eta_{i2}(t), \\ y_{i3} = \displaystyle\int_{t_0}^{t} [f_{i1}(t)\eta_{12}(t) + f_{i01}(t)\eta_{21}(t) + f_{i2}(t)\eta_{11}^2(t)]\,dt \equiv \eta_{i3}(t), \\ y_{im} = \displaystyle\int_{t_0}^{t} [\sum \alpha_{m\mu'} f_{i\mu}(t)[11] \cdots [1m_1] \cdots [\nu 1] \cdots [\nu m_\nu]]\,dt \equiv \eta_{im}(t), \end{cases}$$

where $[pq]$ now denotes $\eta(t)$ with subscripts pq and exponent μ_{pq} and where the c_i are arbitrary parameters which satisfy the inequality

$$(15)\qquad | c_i | \leqq c.$$

Hence a formal solution of the system of differential equations (2) is

$$(16)\qquad x_i = \sum_{m=1}^{\infty} \eta_{im}(t) \qquad (i = 1, 2, \cdots),$$

where the functions $\eta_{im}(t)$ are defined by (14).

2. **Proof of the convergence of the series** (16). In order to obtain a satisfactory dominating function, consider first the differential equation

$$\frac{dX}{dU} = 1 + \sum_{h=1}^{\infty} A_\mu X^{\mu_1+\cdots+\mu_\nu} \qquad (\nu = 1, \cdots, n), \tag{17}$$

where h is defined by (3), n is an arbitrary positive integer and U is the independent variable. As n is a finite positive integer, it is clear that ν is finite, whereas ν takes on every finite positive integral value in (6). It follows from the assumption II that M is an upper bound of the right member of (17), and this right member will converge absolutely when the inequality (7) is satisfied for every positive integral value of n. Hence the right member of (17) can be arranged as a power series in X in the form

$$\frac{dX}{dU} = 1 + \sum_{k=1}^{\infty} B_k X^k, \tag{18}$$

where B_k is the sum of all the A_μ for which

$$\mu_1 + \mu_2 + \cdots + \mu_\nu = k \qquad (\nu = 1, \cdots, n). \tag{19}$$

Hence

$$\begin{cases} B_1 = A_1 + A_{01} + \cdots + A_{0\cdots 01} & (n \text{ terms}), \\ B_k = \sum A_\mu & (\mu_1 + \cdots + \mu_\nu = k;\ k = 1, 2, \cdots). \end{cases} \tag{20}$$

A solution of the differential equation (18) can be found in terms of a power series in U in the form

$$X = U + \frac{1}{2}B_1 U^2 + \left(\frac{1}{3!}B_1^2 + \frac{1}{3}B_2\right)U^3 + \left(\frac{1}{4!}B_1^3 + \frac{1}{3}B_1B_2 + \frac{1}{4}B_3\right)U^4 + \cdots, \tag{21}$$

which will converge when[4]

$$U < \frac{R}{2M}. \tag{22}$$

Consider now the system of differential equations

$$\frac{dX_i}{dU} = 1 + \sum_{h=1}^{\infty} A_\mu X_1^{\mu_1} X_2^{\mu_2} \cdots X_\nu^{\mu_\nu} \qquad (i, \nu = 1, \cdots, n), \tag{23}$$

where h is defined by the relation (3). A formal solution of the system of differential equations (23) can be found by the method of §1 in the form

$$\begin{aligned} X_i = U + \frac{1}{2}A_1U^2 + \left(\frac{1}{3!}A_1^2U^3 + \frac{1}{2}A_{01}U^2 + \frac{1}{3}A_2U^3\right) + \cdots \\ = \sum_{m=1}^{\infty} H_{im}(U) \qquad (i = 1, \cdots, n), \end{aligned} \tag{24}$$

[4] F. R. Moulton, *Differential Equations*, p. 27, formula (12).

where

$$
(25)\qquad \begin{cases} H_{i1}(U) = U, \\ H_{i2}(U) = \tfrac{1}{2}A_1U^2, \\ H_{im}(U) = \displaystyle\int_0^U [\alpha_{m\mu'}A_\mu\{11\}\cdots\{1m_1\}\cdots\{\nu m_\nu\}]\,dU, \end{cases}
$$

where $\{pq\}$ denotes $H(U)$ with subscripts pq and exponent μ_{pq}. Hence

$$
(26)\qquad H_{im}(U) \underset{U}{\equiv} H_{1m}(U) \qquad (i = 1, \cdots, n;\ m = 1, 2, \cdots).
$$

The integrands in the definite integrals of (25) are polynomials in the $H_{ij}(U)$ ($i = 1, \cdots, n; j = 1, \cdots, m - 1$), every term of which is of a degree equal to or greater than one except the $H_{i1}(U)$ which are of the first degree in U. As every term of the integrands in (25) are of at least the first degree in the $H_{ij}(U)$ and the $H_{i1}(U)$ are of the first degree in U, it follows, sequentially, that, after integrating, the $H_{im}(U)$ ($m = 2, 3, \cdots$) will contain no term of the first degree in U. The $H_{im}(U)$ that contain a term of the second degree in U arise from integrands which have a term of the first degree in U. In order that an integrand of (25) contain a term of the first degree in U the equations

$$
(27)\qquad \mu_{11} + \cdots + \mu_{1m_1} + \cdots + \mu_{\nu1} + \cdots + \mu_{\nu m_\nu} = 1
$$

and

$$
(28)\qquad \mu_{12} = \cdots = \mu_{1m_1} = \cdots = \mu_{\nu2} = \cdots = \mu_{\nu m_\nu} = 0
$$

must be satisfied. Equation (28) follows from the fact that the $H_{im}(U)$ ($m = 2, 3, \cdots$) contain no term of the first degree in U. Under these conditions equation (12) reduces to

$$
(29)\qquad \mu_{11} + 2\mu_{21} + \cdots + \nu\mu_{\nu1} = m - 1.
$$

As $\nu \leqq n$, it follows from (27) that there is only a finite number of values of m for which equation (29) can be satisfied. Hence there is only a finite number of the $H_{im}(U)$ that contain a term of the second degree in U. Similarly it can be shown that there are only a finite number of the $H_{im}(U)$ that contain a term of the j-th power of U ($j = 1, 2, \cdots$). Therefore the right members of (24) can be arranged, formally, in a power series in U every coefficient of which is a finite number. It follows from (25), (26) that the series (24), when arranged in power series in U, will satisfy the differential equation (17), and as the power series solution of this differential equation is unique, it is clear that the solution (24) is identically equal to the solution (21). Hence the solution (24) will converge when the inequality (22) is satified.

The right member of the differential equation

$$
(30)\qquad \frac{dZ}{dU} = 1 + \sum_{h=1}^{\infty} A_\mu Z^{\mu_1+\mu_2+\cdots+\mu_\nu} \qquad (\nu = 1, 2, \cdots),
$$

where h is defined by (3), will converge when $|Z| \leqq R$. The right member of (30), being absolutely convergent, can be arranged in a power series in Z in the form

$$\frac{dZ}{dU} = 1 + \sum_{k=1}^{\infty} C_k Z^k, \tag{31}$$

where the C_k are defined by the equations

$$\begin{cases} C_1 = A_1 + A_{01} + A_{001} + \cdots & \text{(an infinite number of terms)}, \\ C_k = \sum A_\mu & (\mu_1 + \mu_2 + \cdots + \mu_\nu = k;\, k, \nu = 1, 2, \cdots). \end{cases} \tag{32}$$

A solution of the differential equation (31) can be found in terms of a power series in U in the form

$$Z = U + \frac{1}{2}C_1^2 U^2 + \left(\frac{1}{3!}C_1^2 + \frac{1}{3}C_2\right)U^3 + \left(\frac{1}{4!}C_1^3 + \frac{1}{3}C_1C_2 + \frac{1}{4}C_3\right)U^4 + \cdots, \tag{33}$$

which will converge when the inequality (22) is satisfied. It follows from (32) and (20) that

$$\lim_{n \to \infty} B_k = C_k \qquad (k = 1, 2, \cdots), \tag{34}$$

and hence

$$\lim_{n \to \infty} X \underset{U}{\equiv} Z, \tag{35}$$

where X is defined by (21) and Z is defined by (33).

As the series (24) and (21) are identically equal in U when the inequality (22) is satisfied, it follows that the relation (35) is true when X is defined by the series (24). But the limit as n approaches infinity of the solution functions (24) is precisely the series obtained by solving the differential equation (31) by the method of §1. This solution is

$$\begin{aligned} Z &= U + \frac{1}{2}A_1U^2 + \left(\frac{1}{3!}A_1^2U^3 + \frac{1}{2}A_{01}U^2 + \frac{1}{3}A_2U^3\right) + \cdots \\ &= \sum_{m=1}^{\infty} N_{1m}(U). \end{aligned} \tag{36}$$

Hence the series (36) will converge when the inequality (22) is satisfied.

A formal solution of the system of differential equations

$$\frac{dZ_i}{dU} = 1 + \sum_{h=1}^{\infty} A_\mu Z_1^{\mu_1} Z_2^{\mu_2} \cdots Z_\nu^{\mu_\nu} \qquad (i, \nu = 1, 2, \cdots), \tag{37}$$

can be found by the method of §1 in the form

$$\begin{aligned} Z_i &= U + \frac{1}{2}A_1U^2 + \left(\frac{1}{3!}A_1^2U^3 + \frac{1}{2}A_{01}U^2 + \frac{1}{3}A_2U^3\right) + \cdots \\ &= \sum_{m=1}^{\infty} N_{im}(U), \end{aligned} \tag{38}$$

where

$$N_{im}(U) \equiv N_{1m}(U) \qquad (i = 1, 2, \cdots). \tag{39}$$

As the right members of (38) are the same as the right members of (36), it is obvious that the series (38) will converge when the inequality (22) is satisfied.

When the variable t varies from t_0 to t, the variable u varies from zero to u, and it follows from the relation (8) that the variable U varies from c, a non-negative number, to U.

The function $N_{im}(U)$ is defined by the definite integral

$$\begin{aligned} N_{im}(U) &= \int_0^U [\sum \alpha_{m\mu'} A_\mu \{11\} \cdots \{\nu m_\nu\}]\, dU \\ &\geqq \int_c^U [\sum \alpha_{m\mu'} A_\mu \{11\} \cdots \{\nu m_\nu\}]\, dU \\ &= \int_0^u [\sum \alpha_{m\mu'} A_\mu \{11\} \cdots \{\nu m\nu\}]\} \theta(u)\, du, \end{aligned} \tag{40}$$

where $\{pq\}$ now denotes $N(U)$ with subscripts pq and exponent μ_{pq}. It follows from (14) and (40) that

$$| \eta_{im}(t) | \leqq N_{im}(U), \tag{41}$$

and hence the formal solution (16) of the system of differential equations (2) will converge when the inequality (22) is satisfied.

3. **Proof of the existence of the derivatives of the solution functions** (16). The function $\eta_{im}(t)$ has the form

$$\begin{aligned} \eta_{im}(t) &= \int_{t_0}^t [\sum \alpha_{m\mu'} f_{i\mu}(t)[11] \cdots [\nu m_\nu]]\, dt \\ &= \int_{t_0}^t G_{im}(t)\, dt, \end{aligned} \tag{42}$$

where

$$G_{im}(t) = \sum \alpha_{m\mu'} f_{i\mu}(t)[11] \cdots [\nu m_\nu] \tag{43}$$

and

$$| G_{im}(t) | \leqq \alpha_{m\mu'} A_\mu \{11\} \cdots \{\nu m_\nu\} \theta(u). \tag{44}$$

The functions $f_{i\mu}(t)$, $\theta_i(t)$, $\theta(u)$, being integrable on the interval (t_0, t), are at most discontinuous at a set of points E of measure zero.[5] The functions $\eta_{im}(t)$ are continuous for all values of t on the interval (t_0, t) except at the set of points E.

[5] E. J. Townsend, *Functions of Real Variables*, p. 212, Theorem IV.

The series (33) is a power series in U and hence has a derivative with respect to U. As the series (36) is identically equal to the series (33), it follows that the series (36) has the derivative

$$\frac{dZ}{dU} = \sum_{m=1}^{\infty} \frac{d}{dU} N_{1m}(U), \tag{45}$$

which is absolutely and uniformly convergent when

$$|U| \leqq \frac{R'}{2M} \qquad (R' < R). \tag{46}$$

Evidently the derivative of Z with respect to u is defined by the series

$$\frac{dZ}{du} = \sum_{m=1}^{\infty} \frac{d}{dU} N_{1m}(U) \cdot \theta(u), \tag{47}$$

for all values of t on the interval (t_0, t) except at the set of points E of measure zero. Inclose the set E in a sequence of intervals (δ_i) the sum of whose lengths is less than ϵ which is arbitrarily small. Delete this sequence of intervals from the interval (t_0, t). It follows from the inequality (44) that the function $G_{im}(t)$ is bounded and continuous on the deleted interval (t_0, t). Hence

$$\frac{d}{dt} \eta_{im}(t) = G_{im}(t), \tag{48}$$

for all values of t on the deleted interval (t_0, t). The right member of (47) dominates the series

$$\sum_{m=1}^{\infty} \frac{d}{dt} \eta_{im}(t) = \sum_{m=1}^{\infty} G_{im}(t), \tag{49}$$

and hence the series (49) converges absolutely and uniformly on the deleted interval (t_0, t). Therefore the functions defined by the series (16) have derivatives at every point on the deleted interval (t_0, t).[6] As ϵ is arbitrarily small, it follows that the functions defined by the series (16) have derivatives at every point on the interval (t_0, t) except at the set of points E of measure zero.

Conclusion. By expressing the solution functions (16) of the system of differential equations (2) in terms of infinite series of definite integrals, the required hypothesis involves the theory of definite integrals and infinite series, but does not require that the right members of the system of differential equations be bounded or continuous on the path of integration. However, the points of discontinuity form at most a set of points of measure zero.

The method of finding the solutions of the present paper is essentially the same as that used by the author in finding solutions of a system of n differential

[6] Townsend, loc. cit., p. 363, Theorem I.

equations[7] except that the arrangement of the terms of (11) is not the same as the corresponding equations (6) of the previous paper.

A slight modification of the method of finding the solutions in terms of power series in a parameter μ, discussed by F. R. Moulton,[8] for a system of n differential equations will give the system (13) of the present paper. However, if the solutions are required to be power series in μ, then the c_i of the present paper must all be taken equal to zero. It is to be noted that the properties of the right members of the differential equations as functions of the independent variable are not the same in this paper as those used by Moulton.

The underlying principle of the present paper is to express the solutions as infinite series in the form

$$x_i = \sum_{h=1}^{\infty} y_{ih} \tag{50}$$

such that when (50) is substituted in (1) the result can be arranged in the form

$$\sum_{h=1}^{\infty} \frac{dy_{ih}}{dt} = \sum_{h=1}^{\infty} \psi_{ih}(t, y_{jk}) \qquad (j, k = 1, \cdots, h-1), \tag{51}$$

where the $\psi_{ih}(t, y_{jk})$ are integrable when the y_{jk} are continuous functions of t. Thus a formal solution is found by integrating the system

$$\frac{dy_{ih}}{dt} = \psi_{ih}(t, y_{jk}) \qquad (i, h = 1, 2, 3, \cdots). \tag{52}$$

The present paper is an example in which the $\psi_{ih}(t, y_{jk})$ are readily found and the convergence of the solutions can be proved.

Solutions of a more general system of differential equations in infinitely many independent variables can be found by a treatment similar to that of §3 of the author's paper cited in footnote 1.

Heidelberg College.

[7] See reference in footnote 1.

[8] Moulton, loc. cit., Chapter III.

RELATIONS BETWEEN CERTAIN CONTINUOUS TRANSFORMATIONS OF SETS

By R. G. Simond

In a paper on arc-preserving transformations[1] G. T. Whyburn has shown that under certain conditions a transformation is a homeomorphism on each cyclic element of a compact locally connected continuum. The principal result is obtained by using a transformation which is arc-preserving and irreducible. After studying Whyburn's theorems, the writer investigated the problem of finding conditions for a homeomorphism on the whole of any compact locally connected continuum and also of finding other conditions for a homeomorphism on the cyclic elements. In working with continua other than cyclic elements, it was found necessary to define some new transformations and to impose conditions on the sets obtained by the transformations. After true arc-preserving and strongly irreducible transformations were defined, it was found that one of Whyburn's proofs could be modified to give a proof of conditions for a homeomorphism on any compact locally connected continuum. The writer has also defined a tree-preserving transformation and has found additional conditions which will make it a homeomorphism in both cases, first on the cyclic elements and then on the whole of any compact locally connected continuum. To do this, true tree-preserving and strongly monotonic transformations have been defined. After these various transformations were studied, it was found that certain relations existed between them, and theorems about some of these relations are also proved in this paper.

The writer wishes to express her appreciation for the help given by Professor W. L. Ayres of the University of Michigan, who suggested the investigation of this problem and gave valuable criticism and advice during the preparation of this paper.

We shall assume throughout that our space is metric and compact. A knowledge of the theory of cyclic elements will also be assumed. Definitions of and theorems about cyclic elements may be found in an expository paper by Kuratowski and Whyburn.[2]

The symbol $\Phi(A)$ will mean a transformation Φ on the set A, where A is a compact locally connected continuum. All of the transformations which are considered are assumed to be single-valued and continuous. For a single-valued transformation continuity is equivalent to the property that a closed set comes

Received March 31, 1938; presented to the American Mathematical Society, April 9, 1937.

[1] G. T. Whyburn, American Journal of Mathematics, vol. 58(1936), pp. 305–312.

[2] C. Kuratowski and G. T. Whyburn, Fundamenta Mathematicae, vol. 16(1930), pp. 305–331.

from a closed set, and, since compactness is assumed, it follows that a closed set goes into a closed set. These facts can be easily verified and have been found useful in the proofs of some of the theorems.

The following types of transformations are used in the proofs of the theorems in this paper: arc-preserving, true arc-preserving, tree-preserving, true tree-preserving, homeomorphic, contracting, irreducible, strongly irreducible, monotonic, and strongly monotonic. Four of these have been studied previously, but the remaining six, viz., the true arc-preserving, tree-preserving, true tree-preserving, contracting, strongly irreducible, and strongly monotonic transformations are defined for the first time in this paper. The definitions of these transformations will now be given and the reasons for introducing the new ones will be stated.

If $\Phi(A) = B$, Φ is said to be *arc-preserving* if the image of every simple arc in A is either a simple arc or a single point in B.[3] A *true arc-preserving* transformation is defined so as to eliminate the possibility of an arc being carried into a point, that is, the image of every non-degenerate arc in A is a non-degenerate arc in B.

To obtain a transformation which is a little more general, a tree-preserving transformation is defined as follows. If $\Phi(A) = B$, Φ will be said to be *tree-preserving* if the image of every subset of A which is a tree is either a non-degenerate tree or a single point in B. As in the case of the arc-preserving transformation, in order to eliminate the possibility of a tree in A being carried into a single point, a *true tree-preserving* transformation is defined as one in which the image of every non-degenerate tree in A is a non-degenerate tree in B.

The definition of a homeomorphic or topological transformation is well known and will be omitted. A new transformation is introduced which only requires that when $\Phi(A) = B$, B shall be homeomorphic with a subset of A. In this case Φ is called a *contracting* transformation.

Whyburn has used the idea of the irreducibility of a transformation in his paper on arc-preserving transformations. Given $\Phi(A) = B$, he defines Φ as an *irreducible* transformation if no proper subcontinuum of A maps onto all of B.[4] In order to extend theorems for cyclically connected sets to the case of any locally connected continuum, it has been found necessary to define a new transformation in which subcontinuum of A is replaced by any closed subset of A. That is, the condition that this subset of A be connected has been removed. Accordingly, given $\Phi(A) = B$, Φ is defined as a *strongly irreducible* transformation if no proper closed subset of A maps onto all of B.

In a paper on the structure of continua,[5] Whyburn states a definition of a monotonic transformation. Given $\Phi(A) = B$, Φ is defined as a *monotonic* transformation if $\Phi^{-1}(b)$ is connected for every $b \in B$. $\Phi^{-1}(b)$ is used to represent

[3] G. T. Whyburn, loc. cit.

[4] G. T. Whyburn, loc. cit.

[5] G. T. Whyburn, Bulletin of the American Mathematical Society, vol. 42(1936), pp. 49–71.

the set of all points x of A such that $\Phi(x) = b$, where $b \epsilon B$. This idea has been used previously by C. B. Morrey.[6] A new definition has been introduced because the condition that $\Phi^{-1}(b)$ be connected was found not to be sufficient in the case of locally connected continua not cyclically connected. A *strongly monotonic* transformation provides not only that $\Phi^{-1}(b)$ be connected, but that it be arcwise connected.

A. **Sufficient conditions for a homeomorphism.** In his paper on arc-preserving transformations[7] Whyburn proved: "If A is a locally connected continuum and $\Phi(A) = B$ is irreducible and arc-preserving, then Φ is a homeomorphism on each true cyclic element of A on which Φ is not constant". In order to obtain conditions which will make Φ a homeomorphism on all of A, the true arc-preserving and strongly irreducible transformations have been defined.

The need of these stronger conditions may be seen from the following examples. First let A be a set which is the sum of three arcs B, C, and D such that B joins c and d, interior points of C and D, respectively. Let Φ be a transformation which is arc-preserving and irreducible on A and such that $\Phi(B) = \Phi(c) = \Phi(d)$ and $\Phi(C) \cdot \Phi(D) = \Phi(B)$. It is evident that Φ is not a homeomorphism on A, but since $\Phi(B)$ is a single point, Φ is not *true arc-preserving*. Next let A be the sum of three arcs B, C, and D such that $B \cdot C = x$, $C \cdot D = y$, and $B \cdot D = 0$. Let Φ be a transformation which is true arc-preserving and irreducible on A and homeomorphic on B, C, and D, and such that $\Phi(B) \cdot \Phi(D) = \Phi(C)$. If c is any interior point of $\Phi(C)$, $\Phi^{-1}(c)$ contains three points of A. Therefore Φ is not a homeomorphism on A. But since $\Phi(A)$ can be obtained from a closed subset of A, namely, $B + D$, Φ is not *strongly irreducible*.

Theorem 1. *If A is a locally connected continuum and Φ is true arc-preserving and strongly irreducible on A, then Φ is a homeomorphism on A.*[8]

Let $\Phi(A) = B$. If Φ is not a homeomorphism, then for some $b \epsilon B$, $\Phi^{-1}(b)$ contains at least two points x and y. Since A is arcwise connected, there is an arc xy in A, and since Φ is true arc-preserving, $\Phi(xy)$ is a true arc. Select a point $p \neq x$ or y so that $\Phi(p)$ is an endpoint of $\Phi(xpy)$. Since A is locally connected, U_x and U_y can be selected as arcwise connected neighborhoods of x and y, respectively, and such that $U_x \cdot U_y = 0$. Since under a single-valued continuous transformation a closed set comes from a closed set, $\Phi^{-1}(p)$ is a closed set. Therefore U_x and U_y can be selected so that $U_x \cdot \Phi^{-1}(p) = 0$ and $U_y \cdot \Phi^{-1}(p) = 0$. Since U_x is open, $A - U_x$ is closed. Since Φ is strongly irreducible, there is a point c in U_x such that $\Phi^{-1}[\Phi(c)] \subset U_x$, for otherwise the closed set $A - U_x$ would map onto all of B. Similarly there is a point d in U_y such that $\Phi^{-1}[\Phi(d)] \subset U_y$. Therefore $\Phi(c) \neq \Phi(d)$. U_x being arcwise connected, there

[6] C. B. Morrey, American Journal of Mathematics, vol. 57 (1935), pp. 17–50.

[7] G. T. Whyburn, loc. cit.

[8] The proof of this theorem is a modification of the proof of the theorem by Whyburn which is stated above.

is an arc from c to x in U_x. Let q be the first point on this arc from c to x which is also on the arc xpy. Similarly let r be the first point on an arc from d to y in U_y which is also on the arc xpy.

Let $M = cq + qp + pr + rd$, which is an arc in A. If Φ is true arc-preserving, $\Phi(M)$ is an arc, and one of the three distinct points $\Phi(c)$, $\Phi(d)$, and $\Phi(p)$ must separate the other two on this arc. $\Phi(qpr)$ is a subarc of $\Phi(xy)$ and this contains a subarc N from $\Phi(q)$ to $\Phi(r)$ which does not contain $\Phi(p)$, for $\Phi(p)$ is an end-point of $\Phi(xy)$ and since $U_x \cdot \Phi^{-1}(p) = 0$ and $U_y \cdot \Phi^{-1}(p) = 0$, $\Phi(q) \neq \Phi(p) \neq \Phi(r)$. Since $\Phi(cq) \cdot \Phi(p) = 0$ and $\Phi(dr) \cdot \Phi(p) = 0$, $\Phi(cq) + N + \Phi(dr)$ is a connected subset of $\Phi(M)$ which does not contain $\Phi(p)$. Therefore $\Phi(p)$ does not separate $\Phi(c)$ and $\Phi(d)$. $\Phi(cq + qp)$ is a connected subset of $\Phi(M)$ which does not contain $\Phi(d)$, and therefore $\Phi(d)$ does not separate $\Phi(p)$ and $\Phi(c)$. Also $\Phi(dr + rp)$ is a connected subset of $\Phi(M)$ which does not contain $\Phi(c)$ and therefore $\Phi(c)$ does not separate $\Phi(p)$ and $\Phi(d)$. Since no one of the three points separates the other two on $\Phi(M)$, $\Phi(M)$ is not an arc. But this is a contradiction, and therefore Φ is a homeomorphism on A.

In order to obtain similar theorems by replacing an arc-preserving by a tree-preserving transformation on A, it was found necessary to place a condition on $\Phi(A)$. If $\Phi(A)$ does not contain a free arc,[9] it is possible to have Φ tree-preserving and irreducible on A and yet not a homeomorphism. This is illustrated by the following example where A is a circle and $\Phi(A) = B$ is a universal tree of order four.

The tree of order four which is described here is slightly different from the one described by K. Menger.[10] Let C be the portion of the (r, θ)-plane which is bounded by $r = 1$. Let I be that portion of C which lies in the first quadrant. Let L be the line $\theta = 45°$ and let I_a and I_b be the two portions of C into which I is divided by L. Let S be a segment of L with $(0, 0°)$ as one endpoint and such that all of its other points lie within I. Select a countable set of points $\{b_n\}$ dense on S and not including the endpoints of S. At b_1 erect two perpendiculars, S_a^1 and S_b^1, such that all of their points except b_1 lie in the interior of D_a^1 and D_b^1, triangles with b_1 as one vertex which lie in I_a and I_b, respectively, and are of diameter $< \frac{1}{2}$. At b_2 erect two perpendiculars, S_a^2 and S_b^2, such that all of their points except b_2 lie in the interior of D_a^2 and D_b^2, triangles with b_2 as one vertex which lie in I_a and I_b, respectively, are of diameter $< \frac{1}{3}$, and are such that $D_\alpha^1 \cdot D_\alpha^2 = 0$ $(\alpha = a, b)$. Continuing in this manner for all the b_n's, in general erect two perpendiculars, S_a^i and S_b^i, at b_i such that all of their points except b_i lie in the interior of D_a^i and D_b^i, triangles with b_i as one vertex which lie in I_a and I_b, respectively, are of diameter $< 1/(i + 1)$, and are such that $D_\alpha^i \cdot \sum_{j=1}^{i-1} D_\alpha^j = 0$. Let $T_{\mathrm{I}}^1 = \sum_{\alpha,i} S_\alpha^i$.

On each S_α^i select a countable set of points dense on the segment and not

[9] An arc xy of a set M is called a free arc if the arc $xy - (x + y)$ is an open subset of M. In other words, no interior point of the arc is a limit point of points of M not on the arc.

[10] K. Menger, *Kurventheorie*, 1932, pp. 318–322.

including its endpoints, and at each of these points erect two perpendiculars, $S_{\alpha c}^{ij}$ and $S_{\alpha d}^{ij}$ (where the superscripts correspond to the numbering in the countable sets of points selected and the subscripts indicate in what portion of I the perpendicular lies), satisfying conditions like those for the perpendiculars to S, so that $T_{\mathrm{I}}^2 = \sum S_{\alpha\beta}^{ij}$ ($\beta = c, d$) will not contain more than a finite number of arcs of diameter $> \epsilon$ for any ϵ.

This process can be continued to obtain $T_{\mathrm{I}}^3, T_{\mathrm{I}}^4, \cdots$ so that there is no free arc in $\bar{B}_{\mathrm{I}}$, where $B_{\mathrm{I}} = S + \sum_{i=1}^{\infty} T_{\mathrm{I}}^i$. Let $\bar{B}_{\mathrm{II}}$, $\bar{B}_{\mathrm{III}}$, and $\bar{B}_{\mathrm{IV}}$ be similar sets constructed in the second, third, and fourth quadrants, respectively. Let $B = \sum_{i=\mathrm{I}}^{\mathrm{IV}} \bar{B}_i$. B is a universal tree of order four.

Let A be the circle $r = 1$, and let A be transformed into B in the following manner. Let Φ_0 be a transformation which carries $(1,0°)$, $(1,90°)$, $(1,180°)$, and $(1,270°)$ into $(0,0°)$ and such that each of the four arcs into which A is divided by these points is carried into a polygon P_i^1 ($i = \mathrm{I}, \cdots, \mathrm{IV}$) lying within the part of C which is in the i-th quadrant, and such that a component of $B - (0,0°)$ is wholly in the interior of it. Let Φ_1 transform four points of P_{I}^1 into b_1 and carry the four arcs into which P_{I}^1 is divided by these points into four polygons $P_{\mathrm{I}i}^2$ ($i = 1, \cdots, 4$) within P_{I}^1 and such that a component of $\bar{B}_{\mathrm{I}} - b_1$ lies wholly within each, and let Φ_1 transform P_{II}^1, P_{III}^1, and P_{IV}^1 in a similar way. Let Φ_2 be a transformation which carries four points of the $P_{\mathrm{I}i}^2$ in which b_2 lies into b_2 and the four arcs into four polygons as above, and let Φ_2 transform a P_{ij}^2 in each of the other quadrants in the same way. If this process is continued for each branch point of B, then $\Phi = \lim \Phi_i$ is a transformation of A into B which is tree-preserving and irreducible, but not a homeomorphism.

THEOREM 2. *If A is a locally connected continuum which is cyclically connected, and Φ is a tree-preserving and irreducible transformation on A, and $\Phi(A)$ contains a free arc, then Φ is a homeomorphism on A.*

Let $\Phi(A) = B$. If Φ is not a homeomorphism on A, then for some $b \in B$, $\Phi^{-1}(b)$ contains at least two points x and y. Let $v \neq b$ be an interior point of a free arc in $\Phi(A)$. Let v_1 be a point of $\Phi^{-1}(v)$. Since A is cyclically connected, there is an arc xv_1y in A. Since Φ is tree-preserving on A, $\Phi(xv_1y)$ is a tree. Let this be T. T contains $b = \Phi(x) = \Phi(y)$ and v and therefore an arc from b to v. Since v is an interior point of a free arc in $\Phi(A)$, there is a subarc pv of the arc bv such that p is an interior point of a free arc in T. Let p_1 be a point of $\Phi^{-1}(p)$ on the arc xv_1y.

Since $p_1 \neq v_1$, p_1 must be on the arc v_1x or the arc v_1y. Suppose p_1 is on the arc v_1y. Select a neighborhood U_{p_1} such that it contains no point of the arc v_1x. Since A being cyclically connected has no cut points, there is a $U_{p_1}^1 \subset U_{p_1}$ such that $A - U_{p_1}^1$ is connected.[11] Since Φ is irreducible on A, there is a point c_1 in $U_{p_1}^1$ such that $\Phi^{-1}[\Phi(c_1)] \subset U_{p_1}^1$, for otherwise $A - U_{p_1}^1$ would map onto all

[11] H. M. Gehman, Proceedings of the National Academy of Sciences, vol. 14(1928), pp. 431-433.

of $\Phi(A)$. Let the diameter of $U^1_{p_1}$ be ϵ and select a sequence of neighborhoods $\{U^i_{p_1}\}$ such that $U^i_{p_1}$ is of diameter $< \epsilon i^{-1}$, and select a sequence $\{c_i\}$ such that $\Phi^{-1}[\Phi(c_i)] \subset U^i_{p_1}$. Since p_1 is a limit point of $\{c_i\}$, p is a limit point of the corresponding sequence $\{c'_i\}$ in $\Phi(A)$. Since p is an interior point of a free arc, there is some c'_k on the arc bv such that $\Phi^{-1}(c'_k) \subset U^k_{p_1}$. But the arc v_1x is a subarc of the arc xy and therefore $\Phi(v_1x)$ is a subtree of T, and since $\Phi(v_1x)$ contains b and v, it also contains c'_k. Since v_1x contains a point of $\Phi^{-1}(c'_k)$, we have a contradiction, and therefore Φ is a homeomorphism on A in this case. The other case may be proved in the same way.

THEOREM 3. *If A is a locally connected continuum, and Φ is a true tree-preserving and strongly irreducible transformation on A, and every arc of $\Phi(A)$ contains a free arc, then Φ is a homeomorphism on A.*

Let $\Phi(A) = B$. If Φ is not a homeomorphism on A, then for some $b \in B$, $\Phi^{-1}(b)$ contains at least two points x and y. Since A is a locally connected continuum, there is an arc xy in A. Since Φ is true tree-preserving, $\Phi(xy)$ is a tree. Let $\Phi(xy) = T$. Let v be any point of T other than b, and let v_1 be a point of $\Phi^{-1}(v)\cdot(\text{arc } xy)$. Let p be some point of the arc bv in T distinct from v and b and on a free arc of $\Phi(A)$. This is possible since every arc of $\Phi(A)$ contains a free arc. Let p_1 be a point of $\Phi^{-1}(p)\cdot(\text{arc } xy)$. Then $p_1 \neq v_1$, for Φ is a single-valued transformation.

Since $p_1 \neq v_1$, p_1 must be on the arc v_1x or the arc v_1y. Suppose p_1 is on the arc v_1y. Let $U^1_{p_1}$ be a neighborhood of p_1 which contains no point of the arc v_1x. Since $U^1_{p_1}$ is open, $A - U^1_{p_1}$ is closed. Since Φ is strongly irreducible on A, there is a point c_1 in $U^1_{p_1}$ such that $\Phi^{-1}[\Phi(c_1)] \subset U^1_{p_1}$, for otherwise $A - U^1_{p_1}$ would map onto all of $\Phi(A)$. Let the diameter of $U^1_{p_1}$ be ϵ and select a sequence of neighborhoods $\{U^i_{p_1}\}$ such that $U^i_{p_1}$ is of diameter $< \epsilon i^{-1}$, and select a sequence $\{c_i\}$ such that $\Phi^{-1}[\Phi(c_i)] \subset U^i_{p_1}$. Since p_1 is a limit point of $\{c_i\}$ in A, p is a limit point of the corresponding sequence $\{c'_i\}$ in $\Phi(A)$. Since p is on a free arc, there is some c'_k on this arc. $\Phi^{-1}(c'_k) \subset U^k_{p_1}$. But the arc v_1x is a connected subset of the arc xy and therefore $\Phi(v_1x)$ is a connected subset of T, and since $\Phi(v_1x)$ contains b and v, it also contains c'_k. Since the arc v_1x contains a point of $\Phi^{-1}(c'_k)$, we have a contradiction, and therefore Φ is a homeomorphism on A.

This theorem without the condition that *every arc* of $\Phi(A)$ contains a free arc is not true. Let A be an arc ab and let c and d be any two distinct interior points of A in the order $acdb$. Let $\Phi(ac)$ and $\Phi(bd)$ be arcs having only $\Phi(c) = \Phi(d)$ in common, and let Φ be a homeomorphism on each of these arcs. Let $\Phi(cd)\cdot\Phi(ac + bd) = \Phi(c)$, where $\Phi(cd)$ is defined in the following manner. Let $\Phi_1(cd) = Y$ be the circle $r = 1$ in the example of the universal tree of order four described above, and let Φ_1 be such that for any $y \in Y$ except $\Phi(c) = \Phi(d)$, $\Phi_1^{-1}(y)$ contains only one point of the arc cd. Let Φ_2 be a transformation which carries Y into a universal tree just as Φ does in the example above. Let $\Phi = \Phi_2\cdot\Phi_1$. Since $\Phi(cd)$ is a universal tree, it contains no free arc. Φ is tree-preserving on A, for since $\Phi(A)$ is a tree, every closed connected subset of $\Phi(A)$

is a tree, and Φ is strongly irreducible on A, for Φ is a homeomorphism on the arc ac and on the arc bd, and all of the arc cd except its endpoints is needed to obtain $\Phi(cd)$. But since $\Phi(A)$ is not an arc, Φ is not a homeomorphism on A.

In order to find other conditions for a homeomorphism on a locally connected continuum, a strongly monotonic transformation has been defined, for it was found necessary to have $\Phi^{-1}(b)$ not only connected but arcwise connected for every $b \in B$.

THEOREM 4. *If A is a locally connected continuum and Φ is a strongly monotonic and true tree-preserving transformation on A, then Φ is a homeomorphism on A.*

Let $\Phi(A) = B$. If Φ is not a homeomorphism on A, then, for some $b \in B$, $\Phi^{-1}(b)$ contains at least two points x and y. Since Φ is strongly monotonic on A, there is an arc xy in A such that $\Phi(xy) = b$. But since Φ is true tree-preserving on A, this is a contradiction, and therefore Φ is a homeomorphism on A.

This theorem is not true if Φ is strongly monotonic but simply tree-preserving instead of true tree-preserving on A. If A is an arc and $\Phi(A)$ is a single point, Φ is strongly monotonic and tree-preserving, but Φ is not a homeomorphism on A.

In Theorem 4, if A is hereditarily locally connected, strongly monotonic can be replaced by monotonic.

THEOREM 5. *If A is a hereditarily locally connected continuum and Φ is a monotonic and true tree-preserving transformation on A, then Φ is a homeomorphism on A.*

Let $\Phi(A) = B$. For any $b \in B$, $\Phi^{-1}(b)$ is closed, for a closed set comes from a closed set under a single-valued continuous transformation. Since Φ is monotonic on A, $\Phi^{-1}(b)$ is connected. Since $\Phi^{-1}(b)$ is closed and connected, it is a subcontinuum of A. Since A is hereditarily locally connected, any subcontinuum is arcwise connected. Therefore Φ is strongly monotonic on A, and by Theorem 4, Φ is a homeomorphism on A.

B. **Other relations between transformations.** In order to prove other relations between transformations, it has been found convenient to establish a few lemmas concerning a configuration which has been called a Y_n-set and to introduce a new subdivision of a tree.

A *Y_n-set* is defined as the sum of n arcs, $n \geqq 3$, having a common endpoint and such that this endpoint is the only point common to any two of the arcs. The arcs will be called the *branches* of the set, and their common endpoint will be called the *center* of the set. Thus $Y_3 = A + B + C$ with center x will mean that A, B, and C are three arcs having x as a common endpoint and $A \cdot B = B \cdot C = A \cdot C = x$.

LEMMA 1. *Given $Y_3 = A + B + C$ with center x. If Φ is arc-preserving on Y_3, then $\Phi(Y_3)$ is an arc or a single point, or $\Phi(Y_3) = \Phi(A) + \Phi(B) + \Phi(C)$, where $\Phi(A)$, $\Phi(B)$, and $\Phi(C)$ are three arcs having $\Phi(x)$ as common endpoint and $\Phi(A) \cdot \Phi(B) = \Phi(B) \cdot \Phi(C) = \Phi(A) \cdot \Phi(C) = \Phi(x)$.*

I. $\Phi(Y_3)$ will be an arc or a single point if $\Phi(A) \subset \Phi(B + C)$. $B + C$ is an arc and therefore since Φ is arc-preserving, $\Phi(B + C)$ is an arc or a single point. But as $\Phi(A) \subset \Phi(B + C)$, $\Phi(Y_3) = \Phi(B + C)$.

II. If $\Phi(Y_3)$ is not an arc or a single point, it follows from I that $\Phi(B)$ is not a subset of $\Phi(C)$ and $\Phi(C)$ is not a subset of $\Phi(B)$, and therefore $\Phi(B)\cdot\Phi(C)$ is a single point, or an arc from some point s to some point t in the arc $\Phi(B + C)$, where s and t are endpoints of the arcs $\Phi(C)$ and $\Phi(B)$, respectively. Since $B \supset x$ and $C \supset x$, $\Phi(B)\cdot\Phi(C) \supset \Phi(x)$.

If $\Phi(B)\cdot\Phi(C)$ is an arc st, let r be the other endpoint of $\Phi(C)$. From I we know that $\Phi(A)$ is not a subset of $\Phi(B + C)$, i.e., there is a point $a \in A$ such that $\Phi(a)$ is not an element of $\Phi(B + C)$. There is an arc ax in A and, since Φ is arc-preserving, $\Phi(ax)$ is an arc in $\Phi(Y_3)$. Let p be the first point that this arc has in common with $\Phi(B + C)$. Since $\Phi(A + C)$ is an arc, p must be an endpoint of $\Phi(C)$. If $p = r$, then $\Phi(C) \subset \Phi(A + B)$ and $\Phi(Y_3)$ is an arc by I. If $p = s$ then $\Phi(A + B)$ is not an arc unless $s = t = \Phi(x)$. Therefore, since Φ is arc-preserving and $\Phi(A + B)$ must be an arc, $\Phi(B)\cdot\Phi(C) = \Phi(x)$ and $\Phi(x)$ is a common end-point of $\Phi(B)$ and $\Phi(C)$.

Similarly it can be shown that $\Phi(A)\cdot\Phi(C) = \Phi(x)$ and $\Phi(A)\cdot\Phi(B) = \Phi(x)$, and $\Phi(x)$ is the common endpoint of any two of the arcs.

LEMMA 2. *Given a Y_n-set with center x, if Φ is arc-preserving on Y_n, then $\Phi(Y_n)$ is an arc or a single point, or $\Phi(Y_n)$ is the sum of the transforms of k of the branches of Y_n ($3 \leqq k \leqq n$) having $\Phi(x)$ as common endpoint and such that $\Phi(x)$ is the only point common to any two of them.*

Let $Y_n = M_1 + M_2 + \cdots + M_n$. If there are two branches of Y_n, M_{i1} and M_{i2}, such that $\Phi(M_{i1} + M_{i2}) = \Phi(Y_n)$, then since $M_{i1} + M_{i2}$ is an arc and Φ is arc-preserving, $\Phi(Y_n)$ is an arc or a single point. If $\Phi(Y_n) = \Phi(M_{i1} + M_{i2} + \cdots + M_{ik})$, where $k \geqq 3$, and no $\Phi(M_{ij}) \subset \Phi(M_{i1} + M_{i2} + \cdots + M_{ik} - M_{ij})$, then by repeated applications of Lemma 1 we may show that $\Phi(x)$ is the common endpoint of $\Phi(M_{i1})$, $\Phi(M_{i2})$, $\cdots$, $\Phi(M_{ik})$ and is the only point common to any two.

COROLLARY 1. *Given $Y_n = M_1 + M_2 + \cdots + M_n$ with center x. If Φ is arc-preserving on Y_n, and $\Phi(Y_n)$ is not an arc (or a single point), then*

(a) *$\Phi(M_i)$ is an arc with $\Phi(x)$ as endpoint [or $\Phi(M_i) = \Phi(x)$]; and*

(b.1) *if $\Phi(M_i)\cdot\Phi(M_j) = \Phi(x)$ for every $M_j \neq M_i$, then $\Phi(M_i)\cdot\Phi(Y_n - N_i) = \Phi(x)$, where $N_i = M_i - x$, or*

(b.2) *if $\Phi(M_i)$ and $\Phi(M_j)$ have points other than $\Phi(x)$ in common, either $\Phi(M_i) \subset \Phi(M_j)$ or $\Phi(M_j) \subset \Phi(M_i)$.*

Remark. Lemma 2 and Corollary 1 are true for x a point of order ω. If the center is a point of order ω, the set will be called a Y_ω-set.

LEMMA 3. *If A is a tree, then $A = \sum_{s=1}^{\infty} A_s$ where*

(1) *A_1 is any arc whose endpoints are endpoints of A;*

(2) *for $s > 1$, $A_s = \sum_{j=1}^{n_s-2} D_{sj}$, where D_{sj} is an arc one endpoint of which is an endpoint of A, and the other endpoint is a point of order n_s and this is the only point which D_{sj} and $\sum_{t=1}^{s-1} A_t$ have in common, and the only point which two of the D_{sj}'s have in common;*

(3) $A - \sum_{s=1}^{\infty} A_s$ *consists entirely of endpoints of A.*

Let $A_1 = L_1$ be some arc joining two endpoints of A. Select the component of $A - L_1$ which has the largest diameter. Since there cannot be more than a finite number of components of diameter $> \delta$ for any δ, the largest one can be selected. Let b_1 be the limit point which this component has on L_1. Let the order of b_1 be n_1. (If b_1 is a point of increasing order, n_1 will be ω, and there will be a countable infinity of components.) There will be $n_1 - 2$ components, C_j, of $A - L_1$ such that $\overline{C_j} \cdot L_1 = b_1$. Since $C_j + b_1$ is closed and bounded, there is a point q such that $\rho(b_1, q) \geqq \rho(b_1, r)$, where r is any other point of C_j. There is an arc D_{2j} with b_1 and an endpoint of A as endpoints and containing q. Select an arc in this manner from each C_j and let the sum of these arcs be A_2; i.e., $A_2 = \sum_{j=1}^{n_1-2} D_{2j}$. Let $L_2 = A_1 + A_2$.

Select the component of $A - L_2$ which has the largest diameter as above, and let b_2 be the limit point which it has on L_2. Let the order of b_2 be n_2. Select an arc in the manner described above from each of the $n_2 - 2$ components of $A - L_2$ having b_2 as a limit point. Let the sum of these arcs be A_3 and let $L_3 = \sum_{s=1}^{3} A_s$. In general b_i is selected so that it is the limit point on L_i of as large a component of $A - L_i$ as possible. A_{i+1} is selected so that it contains one arc, as large as possible, from each component of $A - L_i$ which has b_i as a limit point. $L_j = \sum_{s=1}^{j} A_s$.

Any point p which is not an endpoint of A is an interior point of some arc uv.[12] Let the diameter of the arc up be ϵ. The A_s's were selected in such a way that for any ϵ and for s_1 sufficiently large, no component of $A - \sum_{t=1}^{s_1} A_t$ is of diameter $> \epsilon$. Therefore some point e of the arc up belongs to $\sum_{t=1}^{s_1} A_t$. Similarly some point f of the arc vp belongs to $\sum_{t=1}^{s_2} A_t$ for s_2 large enough. Therefore both e and f belong to $\sum_{t=1}^{s} A_t$ for s equal to the larger of s_1 and s_2. Since there is just one arc from e to f, this is in $\sum_{t=1}^{s} A_t$ and since this arc contains p, p is a point of $\sum_{t=1}^{s} A_t$ for s large enough. This leads to the conclusion that $A - \sum_{s=1}^{\infty} A_s$ consists entirely of endpoints of A.

Since each point of $A - \sum_{s=1}^{\infty} A_s$ is an endpoint and all endpoints are limit points of non-endpoints, $A = \overline{\sum_{s=1}^{\infty} A_s}$.

[12] G. T. Whyburn, Transactions of the American Mathematical Society, vol. 29(1927), p. 385.

THEOREM 6. *If D is a locally connected continuum and Φ is arc-preserving on D, then Φ is tree-preserving on D.*

Although this result is one that might be expected intuitively, we have not been able to find a short method of proof. As the proof is long and detailed, an outline of the method used will be given first. We wish to show that if A is a tree, and $\Phi(A) = B$, and Φ is arc-preserving on A, then B is a tree. This is done by using three monotonic transformations, Φ_1, Φ_2, and Φ_3. We let $\Phi_1(A) = A'$, $\Phi_2(A') = A''$, and $\Phi_3(A'') = A'''$. It is shown that A', A'', and A''' are trees. Φ is defined on A', A'', and A''' in such a way that $\Phi(A') = \Phi(A'') = \Phi(A''') = B$. Finally, A''' is obtained in such a way that Φ is a homeomorphism on A''' and this leads to the desired result that B is a tree. If we use these three monotonic transformations, sets which have certain desired properties are obtained. Φ_1 produces a set A' on which Φ is true arc-preserving. Φ_2 produces a set A'' such that if $z \in B$ has two distinct inverses on A'', they are both on the same free arc. Φ_3 produces a set on which Φ is true arc-preserving and strongly irreducible, that is, as was shown in Theorem 1, a set on which Φ is a homeomorphism.

I

Let A be any subset of D which is a tree and let $\Phi(A) = B$.

Since Φ is not necessarily true arc-preserving on A, there may be arcs $\{\alpha_i\}$ such that $\Phi(\alpha_i)$ is a single point. Let $\{\beta_i\}$ be the set of components of $\sum \alpha_i$. Let Φ_1 be a monotonic transformation which is constant on each β_i, but not on any arc of $A - \sum \beta_i$. Let $\Phi_1(A) = A'$. Since under any monotonic transformation, the image of a tree is a tree,[13] A' is a tree. Let $\Phi_1(\beta_i) = k_i$. Let $\{C_i\}$ be the components of $A - \sum \beta_i$. As Φ_1 is a monotonic transformation, connectedness is an invariant under Φ_1^{-1},[14] and therefore $\Phi_1(C_i) \cdot \Phi_1(C_j) = 0$. If for any point $x \in \Phi_1(C_i)$, $\Phi_1^{-1}(x)$ contains two distinct points x^{-1} and x^{-2}, $\Phi_1^{-1}(x)$ contains the arc from x^{-1} to x^{-2} in A, as Φ_1 is monotonic, but since Φ_1 is not constant on any arc of C_i, this is a contradiction. Therefore $\Phi_1^{-1}(x)$ is a single point of A. If we define Φ on A' so that $\Phi(k_i) = \Phi(\beta_i)$ and $\Phi(x)$ for $x \in \Phi_1 (A - \sum \beta_i)$ so that $\Phi(x) = \Phi[\Phi_1^{-1}(x)]$, then $\Phi(A') = B$ and Φ is true arc-preserving on A'.

II

(A) Let $\{b_i^0\}$ be the branch points of A' such that $\Phi(b_i^0)$ is a point of order one or two in B. Let $\{b_i'\}$ be the branch points of A' such that $\Phi(b_i')$ is a point of order greater than two in B. Let M_{ij} be the arc from b_i' to b_j'. Let $M = \sum M_{ij}$. Since M is a closed subset of A' and A' is a continuous curve, the number of components of $A' - M$ is countable.[15] Let $\{S_{ki}\}$ be the set of these

[13] C. Kuratowski, Fundamenta Mathematicae, vol. 11(1928), p. 182.

[14] G. T. Whyburn, American Journal of Mathematics, vol. 56(1934), p. 295.

[15] R. L. Wilder, Fundamenta Mathematicae, vol. 7(1925), p. 360.

components and let s_k be the limit point of S_{ki} in M. Each s_k must be a branch point of A' or a point of $M - \sum M_{ij}$. Let s_k^0 represent an s_k which is a b_i^0, let s_k' represent an s_k which is a b_i', and let s_k'' represent an s_k which is a point of $M - \sum M_{ij}$.

Case 1. $S_{ki} + s_k^0$. This case is treated below in (B).

Case 2. $S_{ki} + s_k'$. For any $x \epsilon S_{ki}$, by Corollary 1(a), the arc xs_k' goes into an arc with $\Phi(s_k')$ as an endpoint.

Case 3. $S_{ki} + s_k''$. Any s_k'' is a limit point of $\{b_i'\}$ and there is a subsequence of $\{b_i'\}$ which approaches s_k'' along an arc of M. Let $\{b_{ij}'\}$ be such a subsequence which is ordered on the arc. For any $x \epsilon S_{ki}$, by Corollary 1(a), every arc xb_{ij}' goes into an arc with $\Phi(b_{ij}')$ as an endpoint. Since $\Phi(xb_{i,j+1}') \subset \Phi(xb_{ij}')$ and $\lim \Phi\{b_{ij}'\} = \Phi(s_k'')$, $\Phi(xs_k'')$ is an arc having $\Phi(s_k'')$ as an endpoint by the Cantor Product Theorem.

In Cases 2 and 3, let y be any point of S_{ki} distinct from x. Since A' is a tree, the arc xs_k and the arc ys_k have a subarc in common. As Φ is true arc-preserving on A', $\Phi(xs_k)$ and $\Phi(ys_k)$ have points other than $\Phi(s_k)$ in common, and by Corollary 1(b.2), one is a subset of the other. As Φ on $S_{ki} + s_k$ is a continuous transformation on a tree, $\Phi(S_{ki} + s_k)$ is a compact locally connected continuum. Since $\Phi(S_{ki})$ contains no point of order greater than two, $\Phi(S_{ki} + s_k)$ is an arc or a simple closed curve. We next show that it is an arc.

Let E be the set of endpoints of S_{ki}. As E is separable we can let $E = \{x_j\} + Q$, where $\{x_j\}$ is a countable set and every point of Q is a limit point of $\{x_j\}$. If, for any x_j, $\Phi(x_js_k) = \Phi(S_{ki} + s_k)$, then since $\Phi(x_js_k)$ is an arc, $\Phi(S_{ki} + s_k)$ is an arc. In this case let N_i be the arc x_js_k. If $\Phi(x_js_k) \neq \Phi(S_{ki} + s_k)$ for any x_j, then since either $\Phi(x_is_k) \subset \Phi(x_js_k)$ or $\Phi(x_js_k) \subset \Phi(x_is_k)$, there is a subsequence $\{x_j'\}$ of $\{x_j\}$ which can be ordered so that $\Phi(x_j's_k) \subset \Phi(x_{j+1}'s_k)$ and such that for any x_j not in $\{x_j'\}$, $\Phi(x_js_k)$ is contained in $\Phi(x_j's_k)$ for some x_j'.

$$\lim \Phi(x_j's_k) = \lim \sum \Phi(x_j's_k) = \lim \sum \Phi(x_js_k) = \Phi(S_{ki} + s_k).$$

One endpoint of $\Phi(x_j's_k)$ is $\Phi(s_k)$. Let e_j be the other endpoint. As $\{e_j\}$ is an ordered sequence on an arc, it has a sequential limit point p. Let e_j^{-1} be a point of $\Phi^{-1}(e_j)$ on the arc $x_j's_k$ and let $\{e_j^{-1}\}$ be a convergent subsequence of such points. Let $q = \lim \{e_j^{-1}\}$. If $\{e_j^{-1}\}$ converged to s_k, the diameter of the arcs $\{e_j^{-1}s_k\}$ would approach zero and the diameter of the arcs $\Phi\{e_j^{-1}s_k\}$ would approach zero, but $\Phi(e_j^{-1}s_k) = \Phi(x_j's_k)$ and $\Phi(S_{ki} + s_k) = \lim \Phi(x_j's_k)$ cannot be a single point since Φ is true arc-preserving on A'. Therefore $q \neq s_k$. For any e_j^{-1} there is a q_j such that arc $e_j^{-1}s_k$ + arc qs_k = arc q_js_k + arc $q_je_j^{-1}$ + arc q_jq, where no two of these three arcs have more than q_j in common. Since A' is locally connected and $q = \lim \{e_j^{-1}\}$, the limit of the diameters of $\{q_je_j^{-1}\}$ is zero. As q_j is on the arc qs_k, $q = \lim \{q_j\}$.

$$\lim \Phi\{e_j^{-1}s_k\} = \lim \Phi\{e_j^{-1}q_j\} + \lim \Phi\{q_js_k\} = \Phi(qs_k)$$

and

$$\lim \Phi\{e_j^{-1}s_k\} = \lim \Phi\{x_j's_k\} = \Phi(S_{ki} + s_k).$$

But $\Phi(qs_k)$ is an arc and therefore $\Phi(S_{ki} + s_k)$ is an arc. In this case let N_i be the arc qs_k.

Let $K_k = \sum N_i'$ where N_i' is an N_i which has an s_k' as an endpoint and such that $\Phi(N_i')\cdot\Phi(M) = \Phi(s_k')$. By Lemma 2 there is a $K_k' \subset K_k$ where $K_k' = \sum N_i''$ and the N_i'''s are N_i''s such that $\Phi(N_i'')\cdot\Phi(N_j'') = \Phi(s_k')$ and such that $\Phi(K_k') = \Phi(K_k)$. Let H_k be any N_i which has an s_k'' as an endpoint, and let H_k' be any H_k such that $\Phi(H_k')\cdot\Phi(M) = \Phi(s_k'')$.

(B) $A' = M + \sum(S_{ki} + s_k^0) + \sum(S_{ki} + s_k') + \sum(S_{ki} + s_k'')$. Let $V = M + \sum K_k' + \sum H_k'$. For some N_i, $\Phi(S_{ki} + s_k') = \Phi(N_i)$. If this N_i is not in K_k', then either $\Phi(N_i) \subset \Phi(N_j)$, where N_j does belong to K_k', or $\Phi(N_i)$ and $\Phi(M)$ have points other than $\Phi(s_k)$ in common. Let M_{jk} be an M_{ij} such that $\Phi(M_{jk})$ and $\Phi(N_i)$ have points other than $\Phi(s_k)$ in common. By Corollary 1(b.2), either $\Phi(M_{jk}) \subset \Phi(N_i)$ or $\Phi(N_i) \subset \Phi(M_{jk})$. If $\Phi(M_{jk})$ is a proper subset of $\Phi(N_i)$, then $\Phi(b_j')$ is on $\Phi(N_i)$. The point b_j' is the center of a Y_n-set, Y_j. Since $\Phi(b_j')$ is the center of a Y_n-set, $\Phi(Y_j)$, there is a point $x \in \Phi(Y_j) - \Phi(M_{jk})$ and a point $y \in \Phi(N_i) - \Phi(M_{jk})$ such that $\Phi(b_j')$, x, and y do not lie on an arc, but for any $x_1 \in \Phi^{-1}(x)\cdot Y_j$ and $y_1 \in \Phi^{-1}(y)\cdot N_i$, the points x_1, b_j', and y_1 lie on an arc and $\Phi(x_1b_j'y_1)$ is an arc. Therefore in this case $\Phi(N_i) \subset \Phi(M_{jk})$. It follows that $\Phi(S_{ki} + s_k')$ is contained in $\Phi(K_k')$ or in $\Phi(M_{jk})$ for some M_{jk}. The same method of proof can be used to show that $\Phi(S_{ki} + s_k'')$ is contained in $\Phi(H_k')$ or in $\Phi(M_{jk})$ for some M_{jk}. Any s_k^0 which is not an s_k'' on M is an interior point of an M_{ij}. Let W_k be any arc of $S_{ki} + s_k^0$ which has s_k^0 as an endpoint. Since $\Phi(s_k^0)$ is a point of order less than three and Φ is true arc-preserving on A', $\Phi(W_k + M_{ij})$ is an arc. If $\Phi(W_k)$ is not a subset of $\Phi(M_{ij})$, then an arc of $W_k + M_{ij}$ can be found which does not go into an arc, just as in the proof above. Therefore $\Phi(S_{ki} + s_k^0)$ is contained in $\Phi(M_{ij})$. It follows that $\Phi(V) \supset \Phi(A')$, but since V is a subset of A', $\Phi(V) = \Phi(A')$.

Let $\{\gamma_i\}$ be the set of components of $A' - V$ and let g_i be the limit point of γ_i in $A' - V$. Let Φ_2 be a monotonic transformation which is constant on each γ_i, but not on any arc of V. Let $\Phi_2(A') = A''$. Since Φ_2 is monotonic and A' is a tree, A'' is a tree. Let $\Phi_2(\gamma_i) = j_i$. For any point $x \in \Phi_2(V)$, $\Phi_2^{-1}(x)$ is a single point of A' (see proof above that $\Phi_1^{-1}(x)$ is a single point of A). If we define Φ on A'' so that $\Phi(j_i) = \Phi(g_i)$ and $\Phi(x)$ for $x \in \Phi_2(V)$ so that $\Phi(x) = \Phi[\Phi_2^{-1}(x)]$, then $\Phi(A'') = B$.

III

Let X_i represent any branch of a K_k' and let Z_i represent any H_k'. Let $A'' = \sum G_i$ where each G_i is a $\Phi_2(M_{ij})$, a $\Phi_2(X_i)$, or a $\Phi_2(Z_i)$. It can be shown as follows that if G_i contains x but not y and G_j contains y but not x, then $\Phi(x) \neq \Phi(y)$. Let $x_1 = \Phi_2^{-1}(x)\cdot V$ and let $y_1 = \Phi_2^{-1}(y)\cdot V$.

(1) If M_{ij}^1 contains x_1 but not y_1 and M_{ij}^2 contains y_1 but not x_1, the arc from x_1 to y_1 contains a b_k'. Select any b_i' and b_j' such that the arc $b_i'b_k'$ contains x_1 and the arc $b_j'b_k'$ contains y_1. By Corollary 1(a), the images of these two arcs

each have $\Phi(b_k')$ as an endpoint. If they have another point in common, by Corollary 1(b.2), one is a subset of the other. The points b_i' and b_j' are centers of Y_n-sets. Let these be Y_i and Y_j, respectively. Since $\Phi(Y_i)$ and $\Phi(Y_j)$ are Y_n-sets, there is a point $a \epsilon \Phi(Y_i) - \Phi(M)$ and a point $c \epsilon \Phi(Y_j) - \Phi(M)$ such that a, c, and $\Phi(b_k')$ do not lie on an arc, but for any points $a_1 \epsilon \Phi^{-1}(a) \cdot Y_i$ and $c_1 \epsilon \Phi^{-1}(c) \cdot Y_j$, there is an arc $a_1 b_k' c_1$ and therefore $\Phi(a_1 b_k' c_1)$ must be an arc. As this is a contradiction, $\Phi(b_i' b_k')$ and $\Phi(b_j' b_k')$ must have only $\Phi(b_k')$ in common, and $\Phi(x_1) = \Phi(x) \neq \Phi(y_1) = \Phi(y)$.

(2) If M_{ij}^1 contains x_1 but not y_1 and X_i contains y_1 but not x_1, select any b_j' such that the arc from b_j' to b_k', which is an endpoint of X_i, contains x_1. By Corollary 1(a), $\Phi(b_j' b_k')$ and $\Phi(X_i)$ each have $\Phi(b_k')$ as an endpoint. If they have another point in common, by Corollary 1(b.2), one is a subset of the other, but there is an M_{ij}^2 which is a subset of the arc $b_j' b_k'$ and which has b_k' as an endpoint and the X_i's were defined so that $\Phi(X_i) \cdot \Phi(M_{ij}^2) = \Phi(b_k')$. Therefore in this case $\Phi(x_1) \neq \Phi(y_1)$.

(3) If M_{ij}^1 contains x_1 but not y_1 and Z_i contains y_1 but not x_1, the proof that $\Phi(x_1) \neq \Phi(y_1)$ is the same as (2) with Z_i substituted for X_i.

(4) If X_i^1 contains x_1 but not y_1 and X_i^2 contains y_1 but not x_1, let b_i' and b_j' be the b_i''s which are endpoints of X_i^1 and X_i^2, respectively. If $b_i' = b_j'$, $\Phi(X_i^1) \cdot \Phi(X_i^2) = \Phi(b_i')$, for in this case X_i^1 and X_i^2 belong to the same K_k'. Suppose $b_i' \neq b_j'$ and $\Phi(x_1) = \Phi(y_1)$. There is an arc from $\Phi(x_1)$ to $\Phi(b_i')$ in $\Phi(X_i^1)$ and an arc from $\Phi(y_1)$ to $\Phi(b_j')$ in $\Phi(X_i^2)$ and therefore an arc from $\Phi(b_i')$ to $\Phi(b_j')$ in $\Phi(X_i^1) + \Phi(X_i^2)$. There is also an arc from $\Phi(b_i')$ to $\Phi(b_j')$ in $\Phi(M)$. From (2) we know that $\Phi(X_i^1) \cdot \Phi(M) = \Phi(b_i')$ and $\Phi(X_i^2) \cdot \Phi(M) = \Phi(b_j')$. Therefore $\Phi(x_1 b_i') + \Phi(b_i' b_j') + \Phi(b_j' y_1)$ contains a simple closed curve, but the image of the arc $x_1 b_i' b_j' y_1$ must be an arc. Therefore $\Phi(x_1) \neq \Phi(y_1)$.

(5) If Z_i^1 contains x_1 but not y_1 and Z_i^2 contains y_1 but not x_1, the proof that $\Phi(x_1) \neq \Phi(y_1)$ is the same as the second part of (4) ($b_i' \neq b_j'$, for $Z_i^1 \cdot Z_i^2 = 0$) with Z_i substituted for X_i.

(6) If X_i contains x_1 but not y_1 and Z_i contains y_1 but not x_1, the proof that $\Phi(x_1) \neq \Phi(y_1)$ is the same as the second part of (4) (for $b_i' \neq b_j'$) with X_i substituted for X_i^1 and Z_i substituted for X_i^2 and (3) used to show that $\Phi(Z_i) \cdot \Phi(M) = \Phi(b_j')$.

Since these six cases cover all the possibilities, if for any $z \epsilon B$, $\Phi^{-1}(z)$ contains two distinct points in $\sum G_i$, they both belong to the same G_i. Select any δ and any G_i. Let this be G_1. If for any $z_j \epsilon B$ there are two points u_1 and v_1 of $\Phi^{-1}(z_j) \cdot G_1$ such that $\rho(u_1 v_1) > \delta$ (where $\rho(u_1 v_1)$ is the diameter of the arc $u_1 v_1$), let c be any interior point of this arc. Let e_1 and e_2 be the endpoints of G_1, where u_1 is on the arc $e_1 c$ and v_1 is on the arc $c e_2$. Let $\{x_i\}$ be the set of all points on the arc $e_1 u_1$ such that there is a point y_i on the arc $v_1 e_2$ such that $\Phi(x_i) = \Phi(y_i)$. From the continuity of the transformation, $\{x_i\}$ is closed and therefore there is a last point of $\{x_i\}$ on the arc from u_1 to e_1. Let this be u'. Similarly there is a last point of $\{y_i\}$ such that $\Phi(y_i) = \Phi(u')$ on the arc from v_1 to e_2. Let this be v'. Let the arc $u'v'$ be T_1. Let T_1' be T_1 minus its endpoints. If

for any $z_k \epsilon B$ there are two points u_2 and v_2 of $\Phi^{-1}(z_k)\cdot(G_1 - T_1')$ such that $\rho(u_2v_2) > \delta$, select an arc $u_2'v_2' = T_2$ in the same way that T_1 was selected. From the way in which T_1 and T_2 are determined, $T_1\cdot T_2 = 0$. Continuing in this manner, we shall have a finite number of T_i's $> \delta$, for G_1 cannot contain more than a finite number of mutually exclusive arcs of diameter $> \delta$. Next use $\frac{1}{2}\delta$ and G_1 minus the sum of the T_i's previously defined. Continue this process with $\frac{1}{3}\delta$, $\frac{1}{4}\delta$, $\cdots$. Repeat with $G_2 - G_1, \cdots, G_n - \sum_{i=1}^{n-1} G_i, \cdots$. After this has been completed for each G_i, $\sum G_i - \sum T_i'$ will not contain two points u_i and v_i such that $\Phi(u_i) = \Phi(v_i)$.

Let Φ_3 be a monotonic transformation which is constant on each T_i, but not on any arc of $A'' - \sum T_i$. Let $\Phi_3(A'') = A'''$. Since Φ_3 is monotonic and A'' is a tree, A''' is a tree. Let $\Phi_3(T_i) = h_i$. For any point $x \epsilon \Phi_3(A'' - \sum T_i)$, $\Phi_3^{-1}(x)$ is a single point of A'' (see proof above that $\Phi_1^{-1}(x)$ is a single point of A). If we define Φ on A''' so that $\Phi(h_i)$ is the image of the endpoints of T_i and $\Phi(x)$ for $x \epsilon \Phi_3(A'' - \sum T_i)$ so that $\Phi(x) = \Phi[\Phi_3^{-1}(x)]$, we can show as follows that $\Phi(A''') = B$. Let e_1 and e_2 be the endpoints of G_i and let u_iv_i be any arc of G_i such that $\Phi(u_i) = \Phi(v_i)$. Let c be any interior point of the arc u_iv_i. Now, if $\Phi^{-1}(c)\cdot \text{arc } e_1u_i = 0$, $\Phi(c)$ is not on $\Phi(e_1u_i)$ and it must be on $\Phi(v_ie_2)$, and therefore $\Phi^{-1}(c)\cdot \text{arc } v_ie_2 \neq 0$. Therefore $\Phi(A''') = \Phi(A'' - \sum T_i) + \Phi\sum(T_i) = \Phi(A'') = B$.

IV

If, for any point $z \epsilon B$, $\Phi_3^{-1}(z)$ contains two distinct points x and y in A''', it follows from (1)–(6) above that they cannot belong to the images of two different G_i's and from the definition of Φ_3 they cannot both belong to the same $\Phi_3(G_i)$ for any i. Therefore at least one of them must be a point of $\Phi_3(A'' - \sum G_i)$. Suppose x is such a point. Let x^{-1} be a point of $\Phi_3^{-1}(x)\cdot(A'' - \sum G_i)$ and let y^{-1} be a point of $\Phi_3^{-1}(y)$. As $\sum G_i$ contains all the points of $\Phi_2(\sum M_{ij} + \sum K_k' + \sum H_k')$, and all the points of $\overline{\sum K_k'} - \sum K_k'$ and of $\overline{\sum H_k'} - \sum H_k'$ are in M, x^{-1} is a point of $\Phi_2(M - \sum M_{ij})$. Let $\Phi_2^{-1}(x^{-1})\cdot V = x^{-2}$. Since x^{-2} is a limit point of a subsequence $\{b_{ij}'\}$ of $\{b_i'\}$ which approaches x^{-2} along an arc of M, x^{-1} is a limit point of $\Phi_2\{b_{ij}'\}$ which approaches x^{-1} along an arc of $\Phi_2(M)$. As A'' is a tree and x^{-1} is an endpoint of $\Phi_2(M)$, the arc $x^{-1}y^{-1}$ has a subarc in common with this arc of $\Phi_2(M)$ and there is a $\Phi_2(b_i')$ on this subarc. Let $\Phi_2(b_i') = b_i''$. Let I be the arc from x^{-1} to b_i'' and let J be the arc from y^{-1} to b_i''. These arcs have subarcs G_i and G_j, respectively, with b_i'' as an endpoint. If $\Phi(I)\cdot\Phi(J) \neq \Phi(b_i'')$, then by Corollary 1(b.2), either $\Phi(I) \subset \Phi(J)$ or $\Phi(J) \subset \Phi(I)$, but this is impossible for it has been shown above that if $G_i\cdot G_j = b_i''$, then $\Phi(G_i)\cdot\Phi(G_j) = \Phi(b_i'')$. Therefore Φ is a homeomorphism on A''', and since A''' is a tree, B is a tree.

THEOREM 7. *If A is a tree and Φ is true arc-preserving on A, then Φ is contracting on A.*

Let $A' = A$, $\Phi_2(A') = A''$ and $\Phi_3(A'') = A'''$, where Φ_2 and Φ_3 are defined as in Theorem 6. Let $V \subset A'$, Φ on A'', and Φ on A''' also be defined as in Theorem 6.

We know that Φ_2 is a homeomorphism on V and that $\Phi_2(V) = A''$. Therefore A'' is homeomorphic with V. The transformation Φ_3 is such that $\Phi_3(G_i - \sum T_i')$ is an arc. Any two arcs are homeomorphic and therefore there is a transformation ψ such that $\psi(G_i) = \Phi_3(G_i - \sum T_i')$ and such that ψ is a homeomorphism on each G_i. Since $\Phi_3(G_i) \cdot \Phi_3(G_j) = \Phi_3(G_i \cdot G_j)$, ψ is a homeomorphism on $A'' = \sum G_i$. As $\psi(A'') = A'''$, A''' is homeomorphic with A''. As Φ is a homeomorphism on A''', B is homeomorphic with A'''. Since B is homeomorphic with A''', A''' is homeomorphic with A'', and A'' is homeomorphic with $V \subset A' = A$, Φ is contracting on A.

UNIVERSITY OF MICHIGAN.

COVARIANT CONFIGURATIONS RELATED TO ANALYTIC CURVED SURFACES

By P. O. Bell

I. Introduction

In a study of the projective differential properties of a surface a canonical development for the equation of the surface and the geometric determination of the associated reference tetrahedron are of fundamental importance. Both of these problems were solved by Wilczynski.[1] To solve the latter problem he was led to introduce and characterize geometrically the quadric known as the canonical quadric. His method of characterizing this quadric, however, was very complicated. Bompiani[2] has offered a distinctly different characterization, and Stouffer[3] has found a simple method of locating the quadric. Green[4] obtained an expansion which serves to represent a series of canonical developments, including that obtained by Wilczynski. He used, however, Wilczynski's determination of the canonical quadric to characterize a reference tetrahedron. The author[5] presented, in a recent paper, a simple method of completing the determination of the tetrahedron associated with any one of the various canonical developments of Green, without using Wilczynski's quadric. The immediate applications of this method to the theory of surfaces prompted the author to undertake the present investigation.[6]

Let us consider a general curved surface S, referred to its asymptotic net as parametric, with the fundamental differential equations in Wilczynski's canonical form

$$y_{uu} + 2by_v + fy = 0, \qquad y_{vv} + 2a'y_u + gy = 0. \tag{1}$$

Received April 18, 1938; in revised form, June 8, 1938.

[1] E. J. Wilczynski, *Projective geometry of curved surfaces*, (Memoirs 2–3), Transactions of the American Mathematical Society, vol. 9(1908), pp. 79–120; 293–315.

[2] E. Bompiani, *Fascio di quadriche di Darboux e normale proiettiva in un punto di una superficie*, Reale Accademia dei Lincei, Rendiconti, (6), vol. 6(1927), pp. 187–190.

[3] E. B. Stouffer, *A geometrical determination of the canonical quadric of Wilczynski*, Proceedings of the National Academy of Sciences, (18), vol. 3(1932), pp. 252–255.

[4] G. M. Green, *Memoir on the general theory of surfaces and rectilinear congruences*, Transactions of the American Mathematical Society, vol. 20(1919), pp. 79–153.

[5] P. O. Bell, *Tetrahedra associated with canonical expansions for a curved surface*, Bulletin of the American Mathematical Society, vol. 41(1935), pp. 353–355.

[6] The results of this study are presented in complete detail in the author's doctoral dissertation, University of California, 1936. The author takes this opportunity to acknowledge his indebtedness to Miss P. Sperry and to Professor E. B. Stouffer for the encouragement and many helpful suggestions which they offered during the preparation of the present paper.

Using the notation introduced in the celebrated memoir by Green (loc. cit.), let us consider the parametric vector equations

$$(2) \quad y = y(u, v), \quad \rho = y_u - \beta y, \quad \sigma = y_v - \alpha y, \quad \tau = y_{uv} - \alpha y_u - \beta y_v + \alpha\beta y,$$

where α and β are arbitrary analytic functions of u and v. Equations (2) define the general homogeneous coördinates of four points which we denote simply by y, ρ, σ and τ when no possible confusion can arise. If the functions β, α are chosen suitably, the expressions for ρ, σ, and τ become covariants, the points ρ, σ, and τ therefore become covariant points, the coefficients in the associated development become absolute invariants, and the development is said to be a canonical development. The geometric determinations of the covariant points ρ, σ which correspond to various canonical developments are well known, but the problem of the characterization of the associated points τ presents serious difficulties.

The author's characterization (loc. cit.) of a point τ which corresponds to a general selection of points ρ and σ is fundamental in this paper. The characterization may be described briefly as follows. Consider an arbitrary set of covariant points ρ_i, σ_j, τ_{ij} whose general coördinates are given by the forms

$$\rho_i = y_u - \beta_i y, \quad \sigma_j = y_v - \alpha_j y, \quad \tau_{ij} = y_{uv} - \alpha_j y_u - \beta_i y_v + \alpha_j \beta_i y.$$

Let l_{ij} denote the line joining the points ρ_i and σ_j and l'_{ij} denote its reciprocal (i.e., the line in the relation R, as defined by Green, with l_{ij}) with respect to S at a point y. In particular let l_{11} and l'_{11} represent the directrices of Wilczynski. Similarly, the canonical edges of Green will be denoted by l_{22} and l'_{22}. The two points τ_{12} and τ_{21} are characterized geometrically and are used as base points in the following theorem to locate the point τ_{rs} which corresponds to ρ_r and σ_s.

THEOREM I. *The points τ_{ps}, τ_{qr} and the point X^{qr}_{ps} of intersection of the two lines l_{pr} and l_{qs} are collinear.*

Thus the points τ_{rs}, τ_{12} and the point X^{12}_{rs} are collinear, as are also the points τ_{rs}, τ_{21}, and X^{21}_{rs}. Hence τ_{rs} is defined as the intersection of two lines each of which passes through two known points.

II. Reciprocal flat pencils

1. **A sequence of perspectivities connecting reciprocal flat pencils.** We shall call a pair of flat pencils whose lines are reciprocal with respect to a surface S at a point y *reciprocal flat pencils* or simply pencils *in the relation R*. Pencils in the relation R are projective inasmuch as their corresponding lines are reciprocal polar lines with respect to any quadric of the pencil of quadrics of Darboux associated with S at y. We are concerned here with the problem of the geometric determination of a sequence of perspectivities connecting a general pair of flat pencils in the relation R. Let the flat pencils be denoted by l_i and l'_i $(i = 1, 2, \cdots)$. We shall, with Green, call the pencil of lines l_i a pencil of the *first kind*, and the pencil of lines l'_i one of the *second kind*. Let ρ_i and σ_i denote

the intersections of the lines l_i with the tangents to the v = const. and u = const. curves, respectively. Let t_1 denote the line[7] yX_{12}^{21} and t_2 the line yX_{11}^{22}. Let N denote the intersection of t_2 with l_2, and J denote the intersection of the lines $\rho_1\sigma_2$ and $\rho_j N$.

Theorem I shows that the points τ_{11}, τ_{12} and ρ_1 are collinear. The points τ_{22}, τ_{12} and σ_2 are also collinear. Hence the plane determined by the points τ_{12}, τ_{22} and τ_{11} intersects the tangent plane in the line $\rho_1\sigma_2$. The point J therefore lies in this plane and hence the line $J\tau_{12}$ intersects the line $\tau_{11}\tau_{22}$ in a point distinct from τ_{11} and τ_{22} which we denote by M. Let us denote the line yM by m. We shall show that m is identical with the line l_j'. We are now in a position to write the sequence of perspectivities:

$$\begin{aligned}(t_1, l_1, l_2, l_j) \overline{\wedge} (y, \rho_1, \rho_2, \rho_j) &\overset{N}{\overline{\wedge}} (X_{11}^{22}, \rho_1, \sigma_2, J)\\ &\overset{\tau_{12}}{\overline{\wedge}} (X_{11}^{22}, \tau_{11}, \tau_{22}, M) \overline{\wedge} (t_2, l_1', l_2', m).\end{aligned}$$

Hence

$$(t_1, l_1, l_2, l_j) \overline{\wedge} (t_2, l_1', l_2', m).$$

But

$$(t_1, l_1, l_2, l_j) \overline{\wedge} (t_2, l_1', l_2', l_j'),$$

since pencils in the relation R are projective. Therefore

$$m \equiv l_j'.$$

To define geometrically this sequence of perspectivities, it suffices to locate τ_{12} and one point of the set τ_{11}, τ_{22}, M. The remaining points are thereby fixed with reference to these and the given points ρ_i and σ_i. The points τ_{12}, τ_{11} and τ_{22} correspond to known lines $\rho_1\sigma_2$, $\rho_1\sigma_1$ and $\rho_2\sigma_2$ and are therefore easily located by applying the method of Part I. As a result we have a new method of determining the tangent planes to the asymptotic ruled surfaces $R^{(u)}$ and $R^{(v)}$ at points ρ and σ, respectively.

Theorem I may be used in conjunction with the geometrical determination of a single point τ to give still another geometrical characterization of lines which are reciprocal to given lines in the tangent plane. The characterization may be described as follows.

Let τ_{12} be a known point corresponding to a given line $\rho_1\sigma_2$. It is well known that the quadric of Wilczynski may be geometrically characterized as soon as a single point τ, which corresponds to a known line $\rho\sigma$, is located. We shall use the point τ_{12} determined in the author's paper (loc. cit.) for this.

THEOREM 1. *The quadric Q of Wilczynski is the unique quadric which has second order contact with the surface S at y, and which contains the lines $\rho_1\tau_{12}$ and $\sigma_2\tau_{12}$.*

[7] For the sake of brevity we shall denote a line determined by a pair of points by placing in juxtaposition the symbols denoting the points.

It may be easily shown that the equation

$$x_2x_3 - x_1x_4 = 0$$

for the quadric Q of Wilczynski is the same for all coördinate systems referred to reference tetrahedra whose vertices are points whose general coördinates are of the forms y, ρ, σ and τ. Thus all points τ, whose coördinates are of the form

$$\tau = y_{uv} - \alpha y_u - \beta y_v + \alpha\beta y$$

lie on the quadric. Then the point τ_{jj}, where $j \neq 1$, which corresponds to a known line $\rho_j\sigma_j$ lies on the quadric, and by Theorem I is collinear with the known points τ_{12} and X_{12}^{jj}. But the point X_{12}^{jj}, where $j \neq 1$, does not lie on the quadric Q. Hence the line $\tau_{12}X_{12}^{jj}$ intersects the quadric Q in exactly two points which are therefore the points τ_{12} and τ_{jj}. The reciprocal line l_j' is determined by the points y and τ_{jj}. The tangent plane to $R^{(u)}$ at ρ_j is the plane determined by the point ρ_j and the line l_j'. Likewise the tangent plane to $R^{(v)}$ at σ_j is the plane determined by the point σ_j and the line l_j'. The point τ_{jj} is not determined in this way if $j = 1$ or $j = 2$. For in either of these cases X_{12}^{jj}, τ_{12} and τ_{jj} all lie on Q and therefore the line $X_{12}^{jj}\tau_{12}$ is a ruling.

The methods of this section have application to the characterization of the projectivity connecting a very important pair of reciprocal flat pencils, namely, the first and second canonical pencils (Green, loc. cit.).

2. **A geometric characterization of the locus of the points τ which correspond to the covariant lines of the first canonical pencil.** Let us denote by ρ_i and σ_i the intersections of the canonical lines of the first kind with the asymptotic tangents $y\rho$ and $y\sigma$, respectively. Let us consider the flat pencils denoted by $\rho_r\sigma_i$ and $\sigma_s\rho_i$ $(r \neq s)$, whose centers are the fixed points ρ_r and σ_s, respectively. Since the ranges σ_i and ρ_i are perspective from the canonical point, the pencils $\rho_r\sigma_i$ and $\sigma_s\rho_i$ are projective. They are not perspective since there is no self-corresponding line. The pencils, therefore, determine by the points of intersection X_{rs}^{ii} a point conic C_{rs} in the tangent plane to the surface S at y, cutting the asymptotic tangents in the points y, ρ_r and σ_s. The conic C_{rs} is geometrically determined with the location of the lines $\rho_r\sigma_r$ and $\rho_s\sigma_s$. The conic C_{rs} is one of a two-parameter family of such conics $C_{\xi\eta}$ corresponding to the number of arbitrary pairs of lines which may be selected in the pencil. From Theorem I, we see that the points $X_{\xi\eta}^{ii}$ of the conics $C_{\xi\eta}$ are collinear with the points $\tau_{\xi\eta}$ and τ_{ii}. Hence we have

Theorem 2. *The cones of lines determined by the points $\tau_{\xi\eta}$ and the points $X_{\xi\eta}^{ii}$ of the corresponding point conics $C_{\xi\eta}$ determine by their common intersections the points τ_{ii}.*

The points τ_{ii} lie in the canonical plane since they correspond to the canonical lines of the first kind. They therefore determine a point conic (non-degenerate since $\xi \neq \eta$) which we shall call the conic C_1 of the determining canonical cones.

We shall not be concerned here with the residual intersection of these cones, for it does not contain points τ. Projective pencils which determine the conic C_1 may be obtained by projecting the pencils $\rho_r\sigma_i$ and $\sigma_s\rho_i$ on the canonical plane π, from the point τ_{rs}, where the pair r, s is arbitrarily chosen from the pairs ξ, η.

THEOREM 3. *The conics $C_{\xi\eta}$ are the conics which pass through the point y of the surface and are tangent at the points ρ_ξ and σ_η to the canonical lines l_ξ and l_η, respectively, since these lines correspond to the common ray $\rho_\xi\sigma_\eta$ in the respective pencils.*

In particular, the conics of the family $C_{r\eta}$, $r =$ const., are tangent to the line l_r at the point ρ_r and the conics of the family $C_{\xi s}$ are tangent to the line l_s at the point σ_s.

THEOREM 4. *The points $\tau_{r\eta}$ (r = const.) are collinear with the point ρ_r (Theorem* I), *and all lie on the quadric of Wilczynski. Hence the line described by $\tau_{r\eta}$ constitutes a ruling of the quadric. Likewise the points $\tau_{\xi s}$ (s = const.) determine a ruling. As r varies, the ruling $\rho_r\tau_{r\eta}$ generates one regulus of the quadric, and as s varies, the ruling $\sigma_s\tau_{\xi s}$ generates the other regulus of the quadric.*

The plane determined by the lines $\rho_r\tau_{rr}$ and $\sigma_r\tau_{rr}$ is tangent to the quadric of Wilczynski at τ_{rr} since $\rho_r\tau_{rr}$ and $\sigma_r\tau_{rr}$ are rulings. This plane intersects the tangent plane to S at y in the line l_r. Moreover, the conics $C_{r\eta}$ are tangent to l_r at ρ_r. Hence we obtain

THEOREM 5. *The cones of lines $X^{ii}_{r\eta}\tau_{r\eta}$ (r = const.) are mutually tangent along the line determined by their vertices $\tau_{r\eta}$ (r = const.) and their common tangent plane coincides with the plane tangent to the quadric of Wilczynski at the point τ_{rr}.*

An analogous relation holds with respect to the cones of lines $X^{ii}_{\xi s}\tau_{\xi s}$ (s = const.) and the line determined by the points $\tau_{\xi s}$ (s = const.).

The equation of a conic C_{rs} of the family of conics $C_{\xi\eta}$, referred to the points y, ρ_r and σ_s, is easily found to be

$$(\beta_s - \beta_r)x_1x_2 + (\alpha_r - \alpha_s)x_1x_3 + (\alpha_r - \alpha_s)(\beta_s - \beta_r)x_2x_3 = 0,$$

where

$$\rho_s = y_u - \beta_s y, \qquad \rho_r = y_u - \beta_r y,$$

$$\sigma_s = y_v - \alpha_s y, \qquad \sigma_r = y_v - \alpha_r y.$$

Since the conics C_{12} of points X^{ii}_{12} and C_{21} of points X^{ii}_{21} are completely defined geometrically and the points τ_{12} and τ_{21} have been characterized geometrically, we have the following theorem.

THEOREM 6. *The conic C_1 of the points τ_{ii} is characterized geometrically as the mutual intersection of the cones of lines $X^{ii}_{12}\tau_{12}$ and $X^{ii}_{21}\tau_{21}$ with the canonical plane (Theorem* I).

The equation for the conic C_{12} referred to y, ρ_1, and σ_2 is

$$(a'_v b + 2a'b_v)(2ba'_u + a'b_u)x_2x_3 + 4a'b(2ba'_u + a'b_u)x_1x_2 - 4a'b(ba'_v + 2a'b_v)x_1x_3 = 0,$$

and the equation for the conic C_{21} referred to y, ρ_2, and σ_1 is

$$(a_v'b + 2a'b_v)(2ba_u' + a'b_u)x_2x_3 - 4a'b(2ba_u' + a'b_u)x_1x_2 + 4a'b(ba_v' + 2a'b_v)x_1x_3 = 0.$$

III. Certain associated systems of curves in the tangent planes

1. **The sequences C_i, C_i', and C_i^*.** Let C denote a curve, covariantly related to a surface S at y, which lies in the tangent plane to S at y. Let us choose as vertices of the tetrahedron of reference for S four covariant points $y, \rho_1, \sigma_1, \tau_{11}$, which are to be chosen in a manner that will simplify the equation for C as much as possible, but such that the line $\rho_1\sigma_1$ does not coincide with any tangent line of C. A sequence of curves C_i associated covariantly with S at y may then be determined by starting from C as follows. Let a variable tangent line to C be denoted by l. Let the intersections of l with the tangents to the $v = \text{const.}$ and $u = \text{const.}$ asymptotic curves of S be denoted by ρ and σ, respectively. The point τ corresponding to $\rho\sigma$ describes a curve C' as the point of contact of l moves over C. Let C' be projected from the point τ_{11}, which does not lie on C', onto the tangent plane to S at y. Let its projection be denoted by C_1. Taking C_1 and repeating the process gives a curve C_2, likewise C_3 is obtained from C_2, and if we continue the process, C_n is obtained from C_{n-1}.

An associated sequence $C^*, C_1^*, C_2^*, \cdots, C_n^*$ is determined by the intersection of the developable surfaces whose edges of regression are the curves $C', C_1', C_2', \cdots, C_n'$ with the tangent plane to S at y. The correspondence between the points of C_i and C_i^* is one-to-one and continuous. We shall say that the curves of C_i and C_i^* are *in the relation* R^* with respect to S at y.

THEOREM 7. *If C and C^* are any two curves in the relation R^* with respect to S at y, the points of C^* lie on the tangent lines of C whose points of contact are the corresponding points of C.*

To prove this let $\rho\sigma$ denote an arbitrary tangent line to C. Let τ denote the point of C' corresponding to $\rho\sigma$. We must show that the tangent line at τ to the curve C' intersects the line $\rho\sigma$. Now by Part II the curve C' lies on the quadric Q of Wilczynski since it is a locus of points τ whose general coördinates are of the form

$$\tau = y_{uv} - \alpha y_u - \beta y_v + \alpha\beta y.$$

Also in Part II it was shown that the tangent plane to the quadric Q at the point τ intersects the tangent plane to S at y in the line $\rho\sigma$ to which it corresponds. Therefore, since the tangent line to C' at τ lies in the tangent plane to Q at τ, this tangent line intersects the line $\rho\sigma$. This is what we wished to prove.

Let us apply the above methods to the study of a number of important covariant curves. In each case we shall obtain the curves C' and C_1 which are associated with the given curve C in the manner described above. We then investigate the geometric relations which these curves have with important

covariant points and lines of the tangent plane which have been defined by geometers heretofore.

2. **The four-point conics of the projected asymptotics.** Let us consider first a general conic of a pencil of four-point conics having contact of the third order with each other. For definiteness let C_ω denote an arbitrary four-point conic of the projected asymptotics (Green, loc. cit.) which correspond to the v = const. asymptotic curve of S at y. If the points y, ρ_2, σ_2, τ_{22}, whose general coördinates are of the forms

$$y, \quad \rho_2 = y_u - \beta_2 y, \quad \sigma_2 = y_v - \alpha_2 y, \quad \tau_{22} = y_{uv} - \alpha_2 y_u - \beta_2 y_v + \alpha_2\beta_2 y,$$

respectively, are chosen as vertices of the reference tetrahedron, the equation of the conic with which we are concerned is

$$3b^2x_2^2 + 3bx_1x_3 - (b_u + 4b\beta_2)x_2x_3 + \omega x_3^2 = 0, \tag{1}$$

where $2b$ is the coefficient of y_v in the differential equations for S and ω is an arbitrary parameter. By choosing $\rho_2\sigma_2$ to be the first canonical edge of Green we obtain the simple form for the equation:

$$3b^2x_2^2 + 3bx_1x_3 + \omega x_3^2 = 0, \qquad \omega = \text{arbitrary parameter.} \tag{1.1}$$

Let $\omega' = \omega/b$. The equation becomes

$$3bx_2^2 + 3x_1x_3 + \omega' x_3^2 = 0, \qquad \omega' = \text{arbitrary parameter.} \tag{1.2}$$

The parametric equations for the conic $C_{\omega'}$ may be taken in the form

$$\begin{cases} x_1 = -\frac{1}{3}(\omega' + 3bt^2), \\ x_2 = t, \\ x_3 = 1. \end{cases} \tag{2}$$

The equation for a general tangent line l to the conic $C_{\omega'}$ is

$$3x_1 + 6btx_2 + (\omega' - 3bt^2)x_3 = 0.$$

The intersections of this tangent with the tangents to the asymptotic v = const. and u = const. curves, respectively, are the points ρ and σ whose coördinates are given by

$$\rho = y_u - (\beta_2 + 2bt)y, \qquad \sigma = y_v - (\alpha_2 - bt^2)y.$$

The corresponding point τ therefore has general coördinates given by

$$\tau = \tau_{22} - [\tfrac{1}{3}(\omega' - 3bt^2)]\rho_2 - 2bt\sigma_2 + [\tfrac{1}{3}(\omega' - 2bt^2)][2bt]y.$$

Hence a parametric representation of the curve $C'_{\omega'}$ is as follows:

$$\begin{cases} x_1 = \frac{1}{3}(\omega' - 3bt^2)(2bt), \\ x_2 = \frac{1}{3}(3bt^2 - \omega'), \\ x_3 = -2bt, \\ x_4 = 1. \end{cases} \tag{3}$$

The equations of the cone generated by l' as t varies is found, by eliminating t among x_2, x_3, and x_4, to be

$$12bx_2x_4 - 3x_3^2 + 4\omega' bx_4^2 = 0. \tag{4}$$

The projection of $C'_{\omega'}$ onto the tangent plane to S at y from the point τ_{22} is a cubic curve, $\Gamma_{\omega'}$, whose equation is found by eliminating t among x_1, x_2, x_3 of equations (3). The equation is

$$3x_1^2x_3 - 4b\omega' x_2^2x_3 - 12bx_1x_2^2 = 0. \tag{5}$$

When $\omega' = 0$, this cubic degenerates into the first canonical edge of Green and the conic whose equation is

$$x_1x_3 - 4bx_2^2 = 0. \tag{6}$$

Equation (5) shows that the cubic $\Gamma_{\omega'}$ has a node at the point σ_2. The equation for the Hessian of $\Gamma_{\omega'}$ is

$$24b^2\omega' x_1x_2^2 + 3b\omega' x_1^2x_3 + 9bx_1^3 - 4b^2\omega' x_2^2x_3 = 0. \tag{7}$$

The intersections of this with the cubic $\Gamma_{\omega'}$ consist of points of inflection of $\Gamma_{\omega'}$ and the node of $\Gamma_{\omega'}$. The points of inflection are therefore the points whose coördinates are

$$(0, 1, 0), \qquad (-4\omega, 6\sqrt{-\omega}, b), \qquad (4\omega, 6\sqrt{-\omega}, -b),$$

where $\omega = b\omega'$. Clearly the three points of inflection are collinear and determine the line whose equation is

$$bx_1 + 4\omega x_3 = 0. \tag{8}$$

The coördinates of the points of inflection of $\Gamma_{\omega'}$ are such that the truth of the following theorem is immediately apparent:

THEOREM 8. *One of the points of inflection of $\Gamma_{\omega'}$ is the intersection of the first canonical edge of Green with the tangent to the asymptotic $v = $ const. curve, and the remaining two points of inflection of $\Gamma_{\omega'}$ lie on conjugate tangents to S at y.*

The tangents to the cubic $\Gamma_{\omega'}$ at the node (0, 0, 1) have the equations

$$\sqrt{3}x_1 - 2\omega x_2 = 0, \qquad \sqrt{3}x_1 + 2\omega x_2 = 0. \tag{9}$$

Therefore

THEOREM 9. *The nodal tangents of the cubic $\Gamma_{\omega'}$ divide harmonically the tangent, $x_4 = x_3 = 0$, to the $v = $ const. asymptotic curve and the first canonical edge of Green, $x_1 = x_4 = 0$.*

The curve $C^*_{\omega'}$, which is in the relation R^* to $C_{\omega'}$ is a conic whose parametric equations may be taken in the form

$$\begin{cases} x_1 = \tfrac{4}{3}b\omega' - 6b^2t^2, \\ x_2 = 2bt, \\ x_3 = -2b, \\ x_4 = 0. \end{cases} \tag{10}$$

Eliminating t among x_1, x_2, x_3, we obtain the equation for $C^*_{\omega'}$

$$9bx_2^2 - 3x_1x_3 - \omega' x_3^2 = 0. \tag{11}$$

THEOREM 10. *The conics $C_{\omega'}$ and $C^*_{\omega'}$ intersect the tangent, $x_2 = x_4 = 0$, to the asymptotic u = const. curve in the same points, whose coördinates are* $(1, 0, 0, 0)$, $(-\omega, 0, 3, 0)$. *The conics $C_{\omega'}$ and $C^*_{\omega'}$ are both tangent at y to the asymptotic v = const. curve of S.*

Results similar to the above are obtained with the rôles of the asymptotic curves u = const. and v = const. interchanged if we replace the four-point conics which we have used by the four-point conics tangent at y to the u = const. asymptotic curve.

3. **The conics passing through a point y which are tangent to neither asymptotic curve of S at y.** Let us take as a second example a general conic C which passes through the point y of the surface S but which is tangent to neither asymptotic curve of S at y. If we choose y and the two other intersections ρ_1 and σ_1 of the conic with the asymptotic tangents as the three vertices of the reference triangle, the equation of the conic C is of the form

$$x_1x_2 + Ax_1x_3 + Bx_2x_3 = 0. \tag{12}$$

A parametric representation is easily found to be

$$\begin{cases} x_1 = A + Bt, \\ x_2 = t(A + Bt), \\ x_3 = -t. \end{cases} \tag{13}$$

The equation for a tangent line to C at a point having the parameter value t is

$$Bt^2x_1 + Ax_2 + (A + Bt)^2x_3 = 0.$$

The general coördinates of the points of intersection of this tangent with the tangents to the asymptotic u = const. and v = const. curves are given respectively by

$$\sigma = y_v - [\alpha_1 + (A + Bt)^2/Bt^2]y \quad \text{and} \quad \rho = y_u - [\beta_1 + A/Bt^2]y.$$

Hence the points τ which lie on the lines which are the reciprocals with respect to S at y of the tangent lines to C have general coördinates of the form

$$\tau = \tau_{11} - ([A + Bt]^2/Bt^2)\rho_1 - (A/Bt^2)\sigma_1 + (A[A + Bt]^2/B^2t^4)y.$$

A parametric representation for the locus of points τ is, therefore, given by

$$\begin{cases} x_1 = A(A + Bt)^2, \\ x_2 = -Bt^2(A + Bt)^2, \\ x_3 = -ABt^2, \\ x_4 = B^2t^4. \end{cases} \tag{14}$$

The homogeneous equation for the cone determined by y and the locus of points τ corresponding to the tangent lines to C is found, by eliminating t among x_2, x_3, and x_4, to be

$$4ABx_3x_4 + (Ax_3 - Bx_4 - x_2)^2 = 0. \tag{15}$$

The projection of the locus of points τ onto the tangent plane from the point τ_{11}, which corresponds to $\rho_1\sigma_1$, is a curve C_1 whose equation, obtained by eliminating t among x_1, x_2, and x_3, is found to be

$$4ABx_1x_2x_3^2 + (x_1x_2 - Ax_1x_3 + Bx_2x_3)^2 = 0. \tag{16}$$

This is a tricuspidal quartic C_1 having cusps at each vertex of the triangle of reference in the tangent plane to S at y. The equations of the cuspidal tangents at the points y, ρ_1, σ_1 are respectively

$$x_2 - Ax_3 = 0, \qquad x_1 + Bx_3 = 0, \qquad Ax_1 + Bx_2 = 0.$$

THEOREM 11. *The three cuspidal tangents of the quartic C_1 intersect in the point whose local coördinates are* $(-B, A, 1)$.

The tangents to the conic C at the points y, ρ_1, and σ_1 have the respective equations

$$x_2 + Ax_3 = 0, \qquad x_1 + Bx_3 = 0, \qquad Ax_1 + Bx_2 = 0.$$

Hence, we have

THEOREM 12. *If C is any conic lying in the tangent plane to S at y and having as tangent line at y a line distinct from either asymptotic tangent to S, the associated curve C_1 is a tricuspidal quartic curve whose cusps are at the points y, ρ_1, and σ_1, which are the intersections of the conic C with the asymptotic tangents. The cuspidal tangents at ρ_1 and σ_1 coincide with the tangents to C at these points. The cuspidal tangent to C_1 at y and the tangent to C at y are harmonic conjugates with respect to the asymptotic tangents to S at y.*

Let us apply the results proved above to the particular case in which the conic C is the covariant conic C_{12} which we have studied in an earlier section. The equation for this conic is of the form (12), where

$$A = -\frac{ba'_v + 2a'b_v}{a'b_u + 2ba'_u}, \qquad B = \frac{ba'_v + 2a'b_v}{4a'b}.$$

The equation for the associated tricuspidal quartic C_1 is

$$\begin{aligned} 16a'b(a'b_u + 2ba'_u)(ba'_v + 2a'b_v)^2x_1x_2x_3^2 - [4a'b(a'b_u + 2ba'_u)x_1x_2 \\ + 4a'b(ba'_v + 2a'b_v)x_1x_3 + (ba'_v + 2a'b_v)x_2x_3]^2 = 0. \end{aligned} \tag{17}$$

THEOREM 13. *The intersection of the cuspidal tangents to C_1 is the canonical point. The cuspidal tangent to C_1 at y is the first canonical tangent t_1. The cuspidal tangents at the points ρ_1 and σ_2 are, respectively, the first directrix of Wilczynski and the first canonical edge of Green.*

UNIVERSITY OF CALIFORNIA AND UNIVERSITY OF KANSAS.

BOOLEAN FUNCTIONS OF BOUNDED VARIATION

By William D. Duthie

Introduction. The properties of the relation

$$a \Delta b = ab' + a'b$$

suggest its use in framing a definition of a Boolean function of bounded variation similar to that for the real function of a real variable.[1] It is the purpose of this paper to frame such a definition and deduce the restrictions imposed on the function by it. These conditions in turn make possible a re-interpretation of many of the properties of the Δ-relation in terms of this restricted class of functions.

I. Functions of one variable

Definition 1. *The Boolean function $f(x)$ is said to be of bounded variation in the domain* $(0, B)$ *provided the sum*

$$\sum_{\substack{i=1 \\ j=1}}^{n} [f(\alpha_i) \Delta f(\alpha_j)] \tag{1}$$

is different from 1 (*the universal class*) *for all α_i, α_j subject to the conditions*

$$0 < \alpha_i < B, \tag{2}$$

$$\sum_{i=1}^{n} \alpha_i = B, \tag{3}$$

$$\alpha_i \alpha_j = 0 \qquad \text{for } i \neq j. \tag{4}$$

Definition 2. *If the sum* (1) *is null* $(= 0)$, *$f(x)$ is said to be an improper function of bounded variation in the domain* $(0, B)$.

Theorem I. *A necessary and sufficient condition that the function* (*in normal form*)

$$f(x) = ax + bx'$$

be a function of bounded variation in the domain $(0, B)$ *is*

$$(a \Delta b)B \neq 1. \tag{5}$$

Received May 14, 1938; in revised form, June 25, 1938.

[1] For a detailed discussion of the properties of the Δ-relation, see Stone, *Postulates for Boolean algebras and generalized Boolean algebras*, American Journal of Mathematics, vol. 57(1935), pp. 703–732.

Proof.

$$\begin{aligned}
\sum_{i,j=1}^{n} [f(\alpha_i) \,\Delta\, f(\alpha_j)] &= \sum_{i,j=1}^{n} [f(\alpha_i)f'(\alpha_j) + f'(\alpha_i)f(\alpha_j)] \\
&= \sum_{i,j=1}^{n} (a\alpha_i + b\alpha_i')(a'\alpha_j + b'\alpha_j') + \sum_{i,j=1}^{n} (a'\alpha_i + b'\alpha_i')(a\alpha_j + b\alpha_j') \\
&= \sum_{i,j=1}^{n} a'b\alpha_i'\alpha_j + \sum_{i,j=1}^{n} ab'\alpha_i'\alpha_j + \sum_{i,j=1}^{n} a'b\alpha_i\alpha_j' + \sum_{i,j=1}^{n} ab'\alpha_i\alpha_j' \\
&= (ab' + a'b) \sum_{i,j=1}^{n} \alpha_i'\alpha_j + (ab' + a'b) \sum_{i,j=1}^{n} \alpha_i\alpha_j' \\
&= (a \,\Delta\, b) \sum_{i=1}^{n} \alpha_i' \sum_{j=1}^{n} \alpha_j + (a \,\Delta\, b) \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \alpha_j' \\
&= (a \,\Delta\, b)B \sum_{i=1}^{n} \alpha_i' \\
&= (a \,\Delta\, b)B,
\end{aligned}$$

since $\sum_{i=1}^{n} \alpha_i' = \left(\prod_{i=1}^{n} \alpha_i\right)' = 1$ by condition (4) of the definition. This result is clearly independent of the manner in which the α_i, α_j were chosen; hence all the conditions of the definition are fulfilled.

An immediate consequence of the above result is the following

COROLLARY 1. *Any Boolean function whatever is of bounded variation in the domain* $(0, B)$, *if* B *is* $\neq 1$.

Therefore if any restriction is to be imposed on the function itself, the domain of definition must be enlarged to $(0, 1)$. When the domain is $(0, 1)$, the function is said to be of bounded variation *everywhere*.

Hence we have

COROLLARY 2. *A necessary and sufficient condition that the function*

$$f(x) = ax + bx'$$

be of bounded variation everywhere is

$$(a \,\Delta\, b) \neq 1. \tag{5'}$$

The condition (5′) is equivalent to each of conditions $a \neq b'$, $a' \neq b$ separately.

Improper functions of bounded variation. It will be of some interest to determine the most general domain in which the function $f(x) = ax + bx'$ is an improper function of bounded variation. By Theorem I, the most inclusive value of B for which the sum (1) vanishes is

$$B = (a \,\Delta\, b)' = ab + a'b'.$$

If $\sum_{i,j=1}^{n} [f(\alpha_i) \Delta f(\alpha_j)] = 0$, then $[f(\alpha_i) \Delta f(\alpha_j)] = 0$ for any α_i, α_j. Hence $f(\alpha_i) = f(\alpha_j)$ and $f(x)$ remains unchanged for all values of x in the domain $(0, ab + a'b')$. Therefore all the variation of $f(x)$ will occur in the domain $(0, ab' + a'b)$. This domain is called the "effective range" of x by Schmidt, who used a different method to arrive at his results.[2]

The condition that a function be an improper function of bounded variation in the domain $(0, ab' + a'b)$ is from Definition 2

$$(a \Delta b)(ab' + a'b) = a \Delta b = 0 \tag{6}$$

or

$$a = b. \tag{7}$$

The function $f(x)$ then takes the form

$$f(x) = a = b,$$

and all improper functions of bounded variation in the domain $(0, ab' + a'b)$ are constants.

Note. Throughout the remainder of this paper when no domain is specified, it is understood to be the domain $(0, ab' + a'b)$ or $(0, 1)$.

THEOREM II. *The product of any function by an improper function of bounded variation different from 1 is a function of bounded variation.*

Proof. $cf(x) = acx + bcx'$. By Corollary 2, $cf(x)$ will be of bounded variation if $ac \Delta bc \neq 1$. But[3] $ac \Delta bc = (a \Delta b)c$. Hence $ac \Delta bc \neq 1$ if $c \neq 1$.

THEOREM III. *If $f(x)$ is a function of bounded variation, $f(x')$ is also a function of bounded variation.*

Proof. Let $f(x) = ax + bx'$. Then $f(x') = bx + ax'$. But[3] $(a \Delta b) = (b \Delta a)$.

THEOREM IV. *The complement of a function of bounded variation is a function of bounded variation.*

Proof. Let $f'(x) = a'x + b'x'$. Then[3]

$$(a' \Delta b') = (a' \Delta b)' = ((a \Delta b)')' = a \Delta b.$$

A similar theorem may be stated for improper functions of bounded variation. The proof is the same as that of Theorem IV.

Functions not of bounded variation. As in the case of improper functions of bounded variation, it is possible to determine the conditions that a function not be a function of bounded variation.

If the function

$$f(x) = ax + bx'$$

[2] Schmidt, *The theory of functions of one Boolean variable*, Transactions of the American Mathematical Society, vol. 23(1922), pp. 212–222.

[3] For proof of the following relation see Stone, loc. cit.

is not a function of bounded variation, then

$$a \Delta b = 1.$$

Hence $a = b'$ and all functions which are not functions of bounded variation take the form

$$f(x) = ax + a'x'. \tag{8}$$

The function $f(x)$ has the property

$$f'(x) = f(x'). \tag{9}$$

From this relation and Theorem III, we have

THEOREM V. *The complement of a function which is not of bounded variation is a function not of bounded variation.*

An odd property of these functions is the following

THEOREM VI. *The product (sum) of two distinct functions not of bounded variation is a function of bounded variation.*

Proof. Let the functions be

$$f(x) = ax + a'x', \qquad g(x) = bx + b'x'$$

with $a \neq b$. Then

$$f(x)g(x) = abx + a'b'x'.$$

Assume $f(x)g(x)$ is not of bounded variation. Then $ab \Delta a'b' = 1$. Hence $ab = (a'b')' = a + b$ and $a = b$. This contradicts the assumption that $a \neq b$.

By Theorem V, $f'(x)$ and $g'(x)$ are not of bounded variation. Therefore $f'(x)g'(x)$ is a function of bounded variation and

$$f(x) + g(x) = [f'(x)g'(x)]'$$

is a function of bounded variation, by Theorem IV.

Summary. The results of Section I may be summarized as follows.

Let C denote the class of functions of bounded variation, C_1 the class of improper functions of bounded variation, and D the class of functions not of bounded variation. Then $C_1 < C$, $CD = 0$.

If $f(x) \epsilon C$ (C_1, CC_1', D) (ϵ = "is a member of"), then $f'(x) \epsilon C$ (C_1, CC_1', D).

If $f(x) \epsilon C_1$, $f(x) \neq 1$, $g(x)$ arbitrary, then $f(x)g(x) \epsilon C$.

If $f(x) \epsilon D$, $g(x) \epsilon D$, $f(x) \neq g(x)$, then $f(x)g(x) \epsilon C$, and $f(x) + g(x) \epsilon C$.

Sums and products of functions of bounded variation are not necessarily of bounded variation; for example, the sum of the functions

$$f(x) = a'x, \qquad g(x) = ax',$$

where $a \neq 0$ and $a' \neq 0$, is not a function of bounded variation.

II. Notation and a lemma

Before we extend the concept of bounded variation to functions of two variables, an abbreviated notation is introduced.

A Boolean function in normal form

$$f(x_1, x_2, \cdots, x_n) = a_1x_1x_2\cdots x_n + a_2x_1'x_2\cdots x_n + \cdots + a_{2^n}x_1'x_2'\cdots x_n'$$

is determined when its coefficients $a_1, a_2, \cdots, a_{2^n}$ are known. Hence the ordered-set of Boolean elements

$$[a_1, a_2, \cdots, a_{2^n}] \tag{10}$$

may be used to designate the function $f(x_1, x_2, \cdots, x_n)$ *in normal form*.

The following properties of the symbol (10) are simply restatements of well known properties of functions in normal form:

$$[a_1, a_2, \cdots, a_{2^n}]' = [a_1', a_2', \cdots, a_{2^n}'], \tag{11}$$

$$[a_1, a_2, \cdots, a_{2^n}] + [b_1, b_2, \cdots, b_{2^n}] = [a_1 + b_1, a_2 + b_2, \cdots, a_{2^n} + b_{2^n}], \tag{12}$$

$$[a_1, a_2, \cdots, a_{2^n}][b_1, b_2, \cdots, b_{2^n}] = [a_1b_1, a_2b_2, \cdots, a_{2^n}b_{2^n}]. \tag{13}$$

In cases where the same indices are used repeatedly the symbol (10) will be further abbreviated to

$$[a_i] \qquad (i = 1, 2, \cdots, 2^n).$$

Relations (11), (12), and (13) will then be

$$[a_i]' = [a_i'], \tag{11'}$$

$$[a_i] + [b_i] = [a_i + b_i], \tag{12'}$$

$$[a_i][b_i] = [a_ib_i] \qquad (i = 1, 2, \cdots, 2^n). \tag{13'}$$

With the above notation it is now possible to state and prove in a comparatively brief manner the following lemma, which is a generalization of relations (11), (12), (13).

LEMMA. If

$$[a_{ij}] = f_i(x_1, x_2, \cdots, x_n) = a_{i1}x_1x_2\cdots x_n + a_{i2}x_1'x_2\cdots x_n + \cdots + a_{i2^n}x_1'x_2'\cdots x_n' \qquad (i = 1, 2, \cdots, k)$$

is a set of k functions (in normal form) of n variables, then the arbitrary function

$$\phi(f_1, f_2, \cdots, f_k)$$

has the form

$$[\phi(a_{1j}, a_{2j}, \cdots, a_{kj})] \qquad (j = 1, 2, \cdots, 2^n).$$

Proof. Since any arbitrary function ϕ is formed by means of a finite number of operations of the types (11), (12), (13), the j-th coefficient of the function ϕ

will be made up of the same operations on the j-th coefficients of the k functions $[a_{ij}]$.

The notation (10) is also useful in indicating special forms of a function. For example, it is well known that a function (in normal form) of n variables can be considered as a function (in normal form) of one variable whose coefficients are functions (in normal form) of $n - 1$ variables. Such a form of the function is indicated by

$$[[a_1, a_2, \cdots, a_{2^{n-1}}], [a_{2^{n-1}+1}, a_{2^{n-1}+2}, \cdots, a_{2^n}]].$$

Each of the inside brackets may in turn be indicated by a pair of brackets, etc.

III. Functions of two variables

As in the case of real functions, there are several ways in which the concept of bounded variation may be extended to Boolean functions of two variables. The method used in this section was selected because of its comparative brevity.

DEFINITION 3. *A Boolean function of two variables*

$$\begin{aligned}\phi(x, y) &= axy + bx'y + cxy' + dx'y' \\ &= [a, b, c, d]\end{aligned}$$

is said to be of bounded variation with respect to the variable y if the function

$$[[a, b], [c, d]]$$

is of bounded variation for all values of x.

THEOREM VII. *A necessary and sufficient condition that the function* $[a, b, c, d]$ *be of bounded variation with respect to the variable y (x) is that the expression*

$$(a \Delta c) + (b \Delta d) \qquad ((a \Delta b) + (c \Delta d))$$

be different from 1.

Proof. Consider the function in the form

$$[[a, b], [c, d]].$$

Then by Corollary II, Section I,

$$[a, b] \Delta [c, d] \neq 1$$

which is equivalent, by the lemma of Section II, to

$$[(a \Delta c), (b \Delta d)] \neq 1. \tag{14}$$

Since (14) must hold for all values of x, it must hold for $x = a \Delta c$ and condition (14) must be strengthened to

$$(a \Delta c) + (b \Delta d) \neq 1. \tag{15}$$

The sufficiency of the condition is obvious.

By a repetition of the above process, considering the function in the form

$$[[a, c], [b, d]],$$

we get the condition

$$(a \Delta b) + (c \Delta d) \neq 1. \tag{16}$$

DEFINITION 4. *If conditions* (15) *and* (16) *hold simultaneously, the function*

$$[a, b, c, d]$$

is said to be of bounded variation.

Conditions (15) and (16) may be combined into a single sufficient condition

$$a + b + c + d \neq 1.$$

This is not a necessary condition, however.

Analogues of other definitions of Section I may be stated for functions of two variables. Likewise theorems similar to those of Section I may be stated and proved, but the similarities are so marked that these definitions, theorems, and proofs are omitted.

UNIVERSITY OF WASHINGTON.

INTERIOR TRANSFORMATIONS ON CERTAIN CURVES

BY G. T. WHYBURN

In this paper a study will be made of interior transformations as applied to compact metric continua. Also results will be established concerning such transformations defined on certain particular classes of curves, such as dendrites and boundary curves. A single-valued continuous transformation $T(A) = B$ is said to be interior [2][1] provided the image of every open set in A is a set open in B. All continua referred to in this paper are assumed to be compact; and if R is an open set, the boundary of R, i.e., the set $\bar{R} - R$, is designated by $F(R)$.

1. **Conditions for lightness.** A transformation $T(A) = B$ is said to be *light* [5] provided that for no $b \epsilon B$ does $T^{-1}(b)$ contain a non-degenerate continuum. Inasmuch as the property of being light is assumed by Stoïlow [2] for all interior transformations, it is of interest to determine certain classes of continua on which all interior transformations are necessarily light.

(1.1) THEOREM. *If A is a locally connected continuum such that the boundary of every region in A is totally disconnected, then every interior transformation which does not carry A into a single point is light.*

Proof. Suppose, on the contrary, that there exists an interior transformation $T(A) = B$ and a point $p \epsilon B$ such that $T^{-1}(p)$ contains a non-degenerate continuum H. Let R be a component of $B - p$ and let $R_1, R_2, \cdots, R_n$ be the components of $T^{-1}(R)$. Then since T is interior, we have

$$H \subset F(R_1) + F(R_2) + \cdots + F(R_n).$$

Accordingly, for some $i \leqq n$, $F(R_i) \cdot H$ must contain a non-degenerate continuum, contrary to hypothesis.

(1.11) COROLLARY. *If A is either* (i) *a dendrite,* (ii) *a locally connected continuum no cyclic element of which has a continuum of condensation, or* (iii) *a continuum every subcontinuum of which contains uncountably many local separating points of A, then every interior transformation on A is light.*

(1.2) THEOREM. *In order that a continuous transformation $T(A) = B$ be light (where A is compact), it is necessary and sufficient that for any $\epsilon > 0$ a $\delta > 0$ exists such that if X is any continuum in B of diameter $< \delta$, each component of $T^{-1}(X)$ is of diameter $< \epsilon$.*

Proof. The sufficiency is immediate. For if we suppose the condition satisfied and that there is a non-degenerate component H of $T^{-1}(p)$ for some $p \epsilon B$,

Received May 30, 1938.

[1] The numbers in brackets refer to the bibliography at the end of the paper.

taking $\epsilon < \delta(H)$ we can find a continuum X in B with $p \in X$, $\delta(X) < \delta$, whereas the component of $T^{-1}(X)$ containing X is necessarily of diameter $> \epsilon$.

To prove the necessity of the condition, let us suppose on the contrary that there exist a sequence $X_1, X_2, \cdots$ of continua in B with $\delta(X_i) \to 0$ and a sequence $H_1, H_2, \cdots$ of components of $T^{-1}(X_1), T^{-1}(X_2), \cdots$, respectively, with $\delta(H_i) \geqq \epsilon > 0$ for each i. Clearly we may suppose the sequences $[X_i]$ and $[H_i]$ convergent. But this gives $\lim X_i = p \in B$, $\lim (H_i) = H \subset T^{-1}(p)$; and since $\delta(H) \geqq \epsilon$, this is contrary to the hypothesis that T is light.

2. Separating point theorems. We begin with a general theorem on cut points of the image space under an interior transformation.

(2.1) THEOREM. *Suppose $T(A) = B$ is interior, where A is a locally connected continuum. Let p be any point of B and let $Q_1, Q_2, \cdots$ be the components of $B - p$. (The number α of these may be 1, any n, or $\aleph_0$.) If I is any infinite subset of $T^{-1}(p)$, there exist components $R_1, R_2, \cdots$ of $T^{-1}(B - p)$ such that for each i, $T(R_i) = Q_i$ and for each i and j, $F(R_i) \cdot F(R_j) \cdot I$ is infinite. Thus if $\alpha > 1$, there exist a true cyclic element C of A and components $C \cdot R_1, C \cdot R_2, \cdots$ of $C - C \cdot T^{-1}(p)$ such that $F(C \cdot R_i) \cdot F(C \cdot R_j) \cdot I$ is infinite for each i and j.*

Proof. $T^{-1}(Q_1)$ consists [5] of just a finite number of components of $A - T^{-1}(p)$; and since $\overline{T^{-1}(Q_1)} \supset T^{-1}(p) \supset I$, some component, say R_1, of $T^{-1}(Q_1)$ is such that $F(R_1)$ contains an infinite subset I_1 of I. Likewise (if $\alpha > 1$) there exists a component R_2 of $T^{-1}(Q_2)$ [which is a component also of $A - T^{-1}(p)$] such that $F(R_2)$ contains an infinite subset I_2 of I_1. In the same way (if $\alpha > 2$), there is a component R_3 of $T^{-1}(Q_3)$ whose boundary contains an infinite subset I_3 of I_2, and so on. If we continue this process indefinitely, or until a set R_α is found, clearly our theorem is satisfied by the sets $R_1, R_2, \cdots$.

Now if $\alpha > 1$, clearly all points of I_2 are conjugate, since $I_2 \subset F(R_1) \cdot F(R_2)$. Hence, if C is the cyclic element of A containing I_2, the latter part of the theorem is satisfied.

Our theorem yields the following corollaries.

(2.11) *If p is a cut point of B, there exist a finite number of cyclic elements of A whose sum contains $T^{-1}(p)$.*

Proof. Take Q_1 and Q_2 as in the theorem. Let the components of $T^{-1}(Q_1)$ and $T^{-1}(Q_2)$ be $R_1^1, R_2^1, \cdots, R_{n_1}^1$ and $R_1^2, R_2^2, \cdots, R_{n_2}^2$, respectively. Let F_j^i be the set of points $F(R_j^1) \cdot F(R_i^2)$ $(j \leqq n_1,\ i \leqq n_2)$. Then for every set F_j^i there exists a cyclic element C_j^i of A containing F_j^i. Furthermore, since there are only $n_1 n_2$ sets F_j^i, and since $\sum_{i=1}^{n_2} \sum_{j=1}^{n_1} F_j^i = F[T^{-1}(Q_1)] = F[T^{-1}(Q_2)] = T^{-1}(p)$, it follows that $\sum_{i=1}^{n_2} \sum_{j=1}^{n_1} C_j^i \supset T^{-1}(p)$ as required.

(2.12) *If every true cyclic element of A is a linear graph, the inverse of each cut point in B is a finite set.*

(2.13) *If p is a cut point of B such that $T^{-1}(p)$ cuts no true cyclic element of A, then $T^{-1}(p)$ is a finite set.*

(2.2) THEOREM. *Let $T(A) = B$ be interior, where A is a locally connected continuum. If x is a point of B such that $B - x$ has just a finite number of components $R_1, R_2, \cdots, R_k$, and if, for each i, n_i is the number* [5] *of components of $T^{-1}(R_i)$, then $T^{-1}(x)$ contains less than $n = \sum_1^k n_i$ cut points of A.*

Proof. Since any set of n cut points of a connected set A cuts A into at least $n + 1$ mutually separated (open) sets and since no open set can map into x, it follows that if $T^{-1}(x)$ contained as many as n cut points of A, we would have

$$A - T^{-1}(x) = A_1 + A_2 + \cdots + A_{n+1},$$

where the sets A_i are disjoint and open. But clearly this is impossible, since each A_i must contain at least one component of $A - T^{-1}(x)$ and there are only n such components because each one is [5] a component of $T^{-1}(R_i)$ for some i.

(2.21) *If A is locally connected and $T(A) = B$ is interior, and if x is a point of B such that $T^{-1}(x)$ cuts no true cyclic element of A into infinitely many components, then $T^{-1}(x)$ contains at most a finite number of cut points of A.*

For by (2.1) it follows that there can be only a finite number of components of $B - x$. Hence (2.2) applies to give (2.21).

(2.22) *If A (hence also* [5] *B) is a dendrite and $T(A) = B$ is interior, then for each $x \epsilon B$, $T^{-1}(x)$ contains only a finite number of cut points of A. Thus if x is any non-endpoint of B, $T^{-1}(x)$ is a finite set.*

It will be noted that the latter statement is also a consequence of (2.11) or of (2.13).

(2.3) THEOREM. *If A is a compact continuum and $T(A) = B$ is interior, then for each $b \epsilon B$, $T^{-1}(b)$ contains at most a countable number of local separating points of A.*

Proof. Suppose on the contrary that $T^{-1}(b)$ contains an uncountable set G of local separating points of A. Then [4] G contains a point x of degree 2 relative to G. Hence b is of order $\leqq 2$. Accordingly there are at most two components B_1 and B_2 of $B - b$. Now for any $\epsilon > 0$ we can find an ϵ-neighborhood U of x in A such that $F(U)$ consists of just two points of G and $U \cdot G$ is uncountable. Since $T(U) \neq b$, there must exist a quasi-component Q of $A - T^{-1}(b)$ so that $Q \subset U$, since $T^{-1}(b) \supset F(U)$. But by a former theorem ([5], result (1.4)), either $T(Q) = B_1$ or $T(Q) = B_2$; and clearly this is impossible for all $\epsilon > 0$.

It may be observed that (1.1) is also a consequence of (2.3).

(2.4) THEOREM. *If x is a local separating point of a continuum A, if $T(A) = B$ is interior and if x and $y = T(x)$ are of the same finite order*[2] *k in A and B, respectively, then y is a local separating point of B.*

Proof. For ϵ sufficiently small, any ϵ-neighborhood of x or of y must have at least k boundary points. Hence if U is a sufficiently small neighborhood of x in A such that $F(U)$ consists of exactly k points $x_1, x_2, \cdots, x_k$, $V = T(U)$ will be an ϵ-neighborhood of y in B with boundary $F(V) = y_1 + y_2 + \cdots + y_k$,

[2] The term *order of a point* as used in this paper refers to the Menger-Urysohn order.

where $y_i = T(x_i)$. Furthermore, $\bar{U}$ and $\bar{V}$ are continua, and for U sufficiently small, $\bar{U}$ will be the sum of two continua U_1 and U_2 such that $U_1 \cdot U_2 = x$ and $U_1 \cdot F(U) \neq 0 \neq U_2 \cdot F(U)$.

Now since every point y_i of $F(V)$ is a limit point of $B - V$, it follows that, for each i, $x_i = \bar{U} \cdot T^{-1}(y_i)$. Accordingly the transformation $T(\bar{U}) = \bar{V}$ is interior.

Now let us suppose, contrary to what we wish to show, that $\bar{V} - y = R$ is connected. Since this set is open in $\bar{V}$, it follows by a theorem of the author's ([5], result (1.4)) that every quasi-component of $T^{-1}(R)$ maps onto all of R under T. But clearly this is impossible, since U_1 contains a quasi-component Q of $T^{-1}(R)$ and there is at least one integer m such that $x_m \subset U_2$, so that $T^{-1}(y_m) \cdot \bar{Q} = 0$. This contradiction proves our theorem.

(2.41) Corollary. *If every subcontinuum of a continuum A contains uncountably many local separating points of A, the same is true of any interior image of A.*

For let $T(A) = B$ be interior and let K be any subcontinuum of B. Since K contains a subcontinuum irreducible between some two points, we may suppose K is irreducible between two points.[3] Let H be any component of $T^{-1}(K)$. By hypothesis H contains an uncountable set G of local separating points of A. Since all but a countable number of the points of G are [3] of order 2 in A, and since not more than two points of K are of order < 2, it follows that G contains an uncountable subset G_1 such that, for every $x \in G_1$, both x and $T(x)$ are of order 2. Thus by (2.4), any $y \in T(G_1)$ is a local separating point of B; and since, by (2.3), for no $y \in T(G_1)$ is $T^{-1}(y) \cdot G$ uncountable, it follows that $T(G_1)$ is uncountable.

3. **Dendrites.** Let $T(A) = B$ be interior where A is a dendrite. For each $x \in B$ let $k(x)$ be the multiplicity of x, i.e., the number of points in $T^{-1}(x)$. Also, for each integer n, let $B(n)$ denote the set of all $x \in B$ with $k(x) = n$.

Theorem. *For any two points $a, b \in B$ and any point x on the arc ab of B we have*

(i) $k(x) \leqq k(a) + k(b) - 1$.

Proof. Suppose, first, that x is a point of order 2, interior to ab. Then there are just two components R_a and R_b of $B - x$ and these contain a and b, respectively. Now it follows by a former theorem ([5], result (1.4)) that each component of $A - T^{-1}(x)$ maps onto either R_a or R_b. Hence there can be at most $k(a)$ of these components mapping onto R_a and at most $k(b)$ mapping onto R_b. Accordingly there are at most $k(a) + k(b)$ components of $A - T^{-1}(x)$; and since each point of $T^{-1}(x)$ must be a point of order $\geqq 2$ and hence be a cut point of A, it follows by (2.2) that $k(x) \leqq k(a) + k(b) - 1$.

Thus every interior point x of ab of order 2 satisfies (i). Since such points

[3] Actually, since A is hereditarily locally connected and this property is invariant under interior transformations, we could suppose, without loss of generality, that K is a simple arc.

are dense on ab and the points x with $k(x) \leqq$ any integer form a closed set, it follows that (i) holds for any x on ab.

We have the following corollaries.

(3.1) *If $k(a) = k(b) = 1$, then $k(x) = 1$ for each $x \epsilon ab$. Accordingly the set $B(1)$ is a continuum, as is also $T^{-1}[B(1)]$.*

(3.2) *In general, for any m, there exists a dendrite K in B such that*

$$B(m) \subset K \subset B(2m - 1).$$

(3.3) *If $k(a) = 1$ [or $k(b) = 1$], then $k(x) \leqq k(b)$ [or $k(x) \leqq k(a)$]. Thus if $B(1) \neq 0$, then for each m, $\sum_1^m B(i)$ is a dendrite.*

4. **Boundary curves.** A locally connected continuum is called a *boundary curve* provided every one of its true cyclic elements is a simple closed curve. By a *node* of a locally connected continuum M is meant either an endpoint of M or a true cyclic element of M containing at most one cut point of M.

(4.1) THEOREM. *The property of being a boundary curve is invariant under interior transformations.*

Proof. Let A be a boundary curve, let $T(A) = B$ be interior, and let B_0 be any true cyclic element of B. If A_0 is any component of $T^{-1}(B_0)$, then $T(A_0) = B_0$, and on A_0, T is interior [5]. Hence by a theorem of the author's ([6], result (3.2)), any node E_0 of A_0 maps onto all of B_0 under T. Since E_0 is a simple closed curve, B_0 can have at most one point of order $\neq 2$; and hence B_0 is also a simple closed curve.

(4.2) THEOREM. *If A is a locally connected continuum, $T(A) = B$ is interior and B is cyclic, then any simple closed curve node of A maps interiorly onto B.*

Proof. Let E be such a node of A and let p be the cut point of A on E. Order the points of $C \cdot T^{-1}[T(p)]$ cyclically on E, $p = p_1, p_2, \cdots, p_k, p_1$. (These points are finite in number since there can be only a finite number of components of $A - T^{-1}[T(p)]$.) Let q' be any point of $B - T(p)$. Each of the arcs p_ip_{i+1} of E maps onto B and hence contains a point q_i of $T^{-1}(q')$. No arc p_ip_{i+1} of E can contain two points of $T^{-1}(q')$ since each component of $A - T^{-1}(q')$ must contain a point of $T^{-1}[T(p)]$. Thus we have the cyclic order $p_1, q_1, p_2, q_2, \cdots, p_k, q_k, p_1$. It follows that there are just two components R and S of $B - (p' + q')$, where $p' = T(p)$; and p_1q_1 maps topologically onto $R + p' + q'$, q_1p_2 maps topologically onto $S + p' + q'$, p_2q_2 onto $R + p' + q'$, and so on to q_kp_1 which maps topologically onto $S + p' + q'$. Accordingly, $T(E) = B$ is interior.

It will be noted that (4.2) may be made to yield (4.1) as a ready consequence.

5. **1-dimensional images of 2-dimensional pseudo-manifolds.** We consider a compact 2-dimensional pseudo-manifold A, i.e., a set A which can be subdivided so as to form a regularly connected homogeneous 2-dimensional complex every 1-simplex of which is on at most two 2-simplexes [1]. As has been

remarked in an earlier paper ([7], p. 488) it is possible to map a sphere interiorly into any dendrite by sending certain indecomposable continua into points. However, if we require that the inverse of each image point be locally connected, we get the following result.

THEOREM. *If A is a 2-dimensional pseudo-manifold, $T(A) = B$ is interior and for each $b \in B$, $T^{-1}(b)$ is locally connected, then if B is 1-dimensional, it is either an arc or a simple closed curve.*

Proof. Since the 1-dimensional connectivity number of B cannot exceed ([6], (6.4)) that of A, it follows that every cyclic element of B is a linear graph. Thus if Y denotes the set (which we will prove to be vacuous) of all points of B of order > 2, it follows that any $y \in Y$ locally separates B into at least three components. Thus, for any $y \in Y$, there exists a region R in B containing y such that there are at least three components R_1, R_2, R_3 of $R - y$ each having y as a limit point. Since $T^{-1}(y)$ is locally connected, it follows that for each $p \in T^{-1}(y)$ there exist components S_1, S_2, S_3 of $T^{-1}(R_1)$, $T^{-1}(R_2)$, $T^{-1}(R_3)$, respectively, such that p is accessible from each of the sets S_1, S_2, S_3. Now since[4] there can be at most a finite number of points accessible from each set of a given triple S_1, S_2, S_3, and since the total number of such triples of components is finite, it follows that $T^{-1}(y)$ is necessarily a finite set. This being true, it follows at once that every point of $T^{-1}(y)$ is a local separating point of A; and since there are only a finite number of local separating points of A, it follows that Y and $T^{-1}(Y)$ are finite sets and furthermore that $A - T^{-1}(Y)$ is connected. Hence $B - Y$ is connected.

Since every point of $B - Y$ is of order 1 or 2, it follows at once that Y can contain at most two points. But if Y contained two points a and b, then if ab is an arc in B, we would have $ab = B$; and if Y contained a single point y, y would be on a simple closed curve J in B and we would have $J = B$ as y is not a cut point. In either case Y would actually be vacuous. Thus every point of B is of order 1 or 2 and accordingly, B is an arc or a simple closed curve.

BIBLIOGRAPHY

1. P. ALEXANDROFF and H. HOPF, *Topologie*, Berlin, 1935.
2. S. STOÏLOW, *Sur les transformations continues et la topologie des fonctions analytiques*, Annales Scientifiques de l'École Normale Supérieure, vol. 63(1928), pp. 347–382.
3. G. T. WHYBURN, *Local separating points of continua*, Monatshefte für Mathematik und Physik, vol. 36(1929), pp. 305–314.
4. G. T. WHYBURN, *Concerning continua of finite degree and local separating points*, American Journal of Mathematics, vol. 57(1935), pp. 11–16.
5. G. T. WHYBURN, *Interior transformations on compact sets*, this Journal, vol. 3(1937), pp. 370–381.
6. G. T. WHYBURN, *The mapping of Betti groups under interior transformations*, this Journal, vol. 4(1938), pp. 1–8.
7. G. T. WHYBURN, *Interior transformations on surfaces*, American Journal of Mathematics, vol. 60(1938), pp. 477–490.

UNIVERSITY OF VIRGINIA.

[4] This follows readily from known plane accessibility theorems. See, for example, Bulletin of the American Mathematical Society, vol. 34(1928), p. 504.

INTEGRAL INEQUALITIES CONNECTED WITH DIFFERENTIAL OPERATORS

By I. Halperin and H. R. Pitt

1. **Introduction.** Suppose that the measurable functions $f(x)$, $f_r(x)$, $q_r(x)$, $q_{r,j}(x)$, $p_r(x)$, $c(x)$ are defined in the finite or infinite interval $\alpha < x < \alpha + a$, that $f_0(x), f_1(x), \cdots, f_{n-1}(x)$ are absolutely continuous in every closed subinterval and that almost everywhere[1] in $(\alpha, \alpha + a)$

$$
\begin{aligned}
&f_0(x) = q_0(x)f(x), \\
(1.1)\qquad &f_{r+1}(x) = f_r'(x) + \sum_{j=0}^{r} q_{r,j}(x)f_j(x) + q_{r+1}(x)f(x) \qquad (r = 0, 1, \cdots, n-1),
\end{aligned}
$$

$$
(1.2)\qquad Tf(x) = \sum_{r=0}^{n} p_r(x)f_r(x) + c(x)f(x),
$$

$$
\begin{aligned}
&|q_{r,j}(x)| \leqq M_1 \qquad && (j = 0, 1, \cdots, r; r = 0, 1, \cdots, n-1), \\
(1.3)\qquad &|q_r(x)| \leqq M_1, \qquad |p_r(x)| \leqq M_2, && (r = 0, 1, \cdots, n), \\
&|c(x)| \leqq M_3,
\end{aligned}
$$

$$
(1.4)\qquad |q_0(x)| \geqq \epsilon > 0, \qquad |p_n(x)| \geqq \epsilon > 0.
$$

We write

$$
(1.5)\qquad A_r = \left[\int_\alpha^{\alpha+a} |f_r(x)|^p\,dx\right]^{1/p}, \qquad B = \left[\int_\alpha^{\alpha+a} |Tf(x)|^p\,dx\right]^{1/p},
$$

where $p \geqq 1$.

Our object in this paper is to prove inequalities of the form

$$
(1.6)\qquad A_r \leqq F(A_0, A_n), \qquad A_r \leqq F(A_0, B),
$$

and to use them to prove certain results about differential operators.

2. **Integral inequalities.**

(2.1) Theorem 1. *Suppose that* (1.1), (1.2), (1.3), *and* (1.4) *are satisfied and that* $\eta > 0$. *Then*

$$
(2.1.1)\qquad A_r \leqq \eta B + KA_0 \qquad (r = 0, 1, \cdots, n-1),
$$

Received June 10, 1938.

[1] The words "almost everywhere" will be understood in similar statements that occur in what follows, although they will be frequently omitted.

where $K = K(\epsilon, M_1, M_2, M_3, n, \eta, a) < \infty$ *is a decreasing function of* a *for fixed* $\epsilon, M_1, M_2, M_3, n, \eta$. *In particular, if* $p_n(x) = 1, p_r(x) = 0$ $(r = 0, 1, \cdots, n - 1)$, *then*

$$A_r \leqq \eta A_n + KA_0 \qquad (r = 0, 1, \cdots, n-1). \tag{2.1.2}$$

First we observe that by replacing $f(x)$ by $q_0(x)f(x)$ and making suitable changes in $p_r(x)$, $q_{r,j}(x)$, we may assume that $q_0(x) = p_n(x) = 1$, $c(x) = 0$, and $q_r(x) = 0$ $(r = 1, 2, \cdots, n)$. Secondly, it is sufficient to establish the inequalities (2.1.1), (2.1.2) for the closed subintervals interior to $(\alpha, \alpha + a)$ (for which the A_r's are clearly finite), since the general validity of (2.1.1), (2.1.2) will follow by a straightforward limit process. Thus we may suppose that $a < \infty$ and that $\alpha = 0$.

We now prove two elementary lemmas.

(2.2) LEMMA 1. *Let* $h > 0$ *and suppose that* $f(x)$ *belongs to* $L^p(a, b + h)$. *Then*

$$\int_a^b \left| \int_0^h f(x+t)\,dt \right|^p dx \leqq h^p \int_a^{b+h} |f(x)|^p\,dx, \tag{2.2.1}$$

$$\int_a^b \left| \int_0^h (h-t)f(x+t)\,dt \right|^p dx \leqq \frac{h^{2p}}{2^p} \int_a^{b+h} |f(x)|^p\,dx. \tag{2.2.2}$$

These inequalities are special cases of a well-known theorem of W. H. Young.[2]

(2.3) LEMMA 2. *If* $0 < h < \frac{1}{4}a$, *then*

$$A_r\left[\frac{a-4h}{a-2h} - \tfrac{1}{2}M_1 h\right] \leqq \frac{2}{h}A_{r-1} + \frac{h}{2}A_{r+1} + M_1(1 + \tfrac{1}{2}h)\sum_{j=0}^{r-1} A_j. \tag{2.3.1}$$

Let

$$g_r(x) = \sum_{j=0}^{r-1} q_{r-1,j}(x)f_j(x) \qquad (r = 1, 2, \cdots, n),$$

so that

$$f_r(x) = f'_{r-1}(x) + g_r(x). \tag{2.3.2}$$

Also, by Minkowski's inequality

$$\left[\int_0^a |g_r(x)|^p\,dx\right]^{1/p} \leqq M_1 \sum_{j=0}^{r-1} A_j. \tag{2.3.3}$$

Now let $0 < h \leqq \xi \leqq a - h$. Let

$$\begin{aligned} \gamma(x) &= \int_0^h f_r(x+t)\,dt && (0 \leqq x < \xi - h), \\ &= 0 && (\xi - h \leqq x < \xi + h), \\ &= \int_0^h f_r(x-t)\,dt && (\xi + h \leqq x \leqq a). \end{aligned}$$

[2] See A. Zygmund, *Trigonometrical Series*, p. 71, 4.16.

Then if $0 \leqq x < \xi - h$,

$$hf_r(x) = \int_0^h f_r(x+t)\,dt - \int_0^h [f_r(x+t) - f_r(x)]\,dt$$

$$= \gamma(x) - \int_0^h (h-t)f_r'(x+t)\,dt.$$

Similarly, if $\xi + h \leqq x < a$,

$$hf_r(x) = \gamma(x) + \int_0^h (h-t)f_r'(x-t)\,dt.$$

If we use Lemma 1, it follows that

$$\int_0^{\xi-h} + \int_{\xi+h}^a |hf_r(x) - \gamma(x)|^p\,dx$$

$$= \int_0^{\xi-h}\left|\int_0^h (h-t)f_r'(x+t)\,dt\right|^p dx + \int_{\xi+h}^a\left|\int_0^h (h-t)f_r'(x-t)\,dt\right|^p dx \tag{2.3.4}$$

$$\leqq \frac{h^{2p}}{2^p}\int_0^\xi |f_r'(x)|^p\,dx + \frac{h^{2p}}{2^p}\int_\xi^a |f_r'(x)|^p\,dx = \frac{h^{2p}}{2^p}\int_0^a |f_r'(x)|^p\,dx.$$

Now by (2.3.2)

$$\int_0^h f_r(x \pm t)\,dt = \int_0^h [f_{r-1}'(x \pm t) + g_r(x \pm t)]\,dt$$

$$= \pm[f_{r-1}(x \pm h) - f_{r-1}(x)] + \int_0^h g_r(x \pm t)\,dt.$$

It follows easily from this and (2.3.3), by use of Minkowski's inequality, that

$$\left[\int_0^{\xi-h} + \int_{\xi+h}^a |\gamma(x)|^p\,dx\right]^{1/p} \leqq 2A_{r-1} + hM_1\sum_{j=0}^{r-1} A_j. \tag{2.3.5}$$

Also, by (2.3.2), (2.3.3) and Minkowski's inequality,

$$\left[\int_0^a |f_r'(x)|^p dx\right]^{1/p} \leqq A_{r+1} + M_1\sum_{j=0}^{r} A_j. \tag{2.3.6}$$

If we combine (2.3.4), (2.3.5), (2.3.6) and use Minkowski's inequality again, we obtain

$$\left[\int_0^{\xi-h} + \int_{\xi+h}^a |hf_r(x)|^p\,dx\right]^{1/p} \leqq 2A_{r-1} + hM_1\sum_{j=0}^{r-1} A_j + \frac{h^2}{2}\left[A_{r+1} + M_1\sum_{j=0}^{r} A_j\right].$$

Hence

$$A_r^p \leqq \left[\frac{2}{h}A_{r-1} + \frac{h}{2}A_{r+1} + \frac{hM_1}{2}A_r + M_1\left(1+\frac{h}{2}\right)\sum_{j=0}^{r-1} A_j\right]^p + \int_{\xi-h}^{\xi+h} |f_r(x)|^p\,dx,$$

and if we integrate this inequality over the range $h \leqq \xi \leqq a - h$ and use Lemma 1, we get

$$(a - 2h)A_r^p \leqq \left[\frac{2}{h}A_{r-1} + \frac{h}{2}A_{r+1} + \frac{hM_1}{2}A_r + M_1\left(1 + \frac{h}{2}\right)\sum_{j=0}^{r-1} A_j\right]^p (a - 2h) + 2hA_r^p.$$

Hence

$$A_r\left[\left(\frac{a - 4h}{a - 2h}\right)^{1/p} - \frac{hM_1}{2}\right] \leqq \frac{2}{h}A_{r-1} + \frac{h}{2}A_{r+1} + M_1\left(1 + \frac{h}{2}\right)\sum_{j=0}^{r-1} A_j,$$

an inequality which is stronger than (2.3.1) since $a - 4h < a - 2h$.

(2.4) *Proof of* (2.1.2). As we use different values of η, it is convenient to write $K(\eta)$ for any function which depends only on ϵ, M_1, M_2, n, η, a and is a decreasing function of a for fixed ϵ, M_1, M_2, n, η. The proof is by induction. Suppose that for any $\eta' > 0$ and $r = 0, 1, \cdots, n - 1$ we have

$$A_r \leqq \eta' A_n + K(\eta')A_0. \tag{2.4.1}$$

It follows from Lemma 2 that

$$A_n \leqq \eta'' A_{n+1} + K(\eta'')\sum_{j=0}^{n-1} A_j. \tag{2.4.2}$$

If we combine (2.4.1) and (2.4.2), we obtain

$$A_n \leqq \eta'' A_{n+1} + nK(\eta'')[\eta' A_n + K(\eta')A_0],$$

$$A_n[1 - n\eta' K(\eta'')] \leqq \eta'' A_{n+1} + nK(\eta'')K(\eta')A_0,$$

and when η', η'' are suitably defined, this can be written

$$A_n \leqq \eta A_{n+1} + K(\eta)A_0.$$

It follows from this and (2.4.1) that

$$A_r \leqq \eta A_{n+1} + K(\eta)A_0 \qquad (r = 0, 1, \cdots, n). \tag{2.4.3}$$

(2.5) *Proof of* (2.1.1). From (1.2) and (1.5) we derive by means of Minkowski's inequality

$$A_n \leqq B + M_2 \sum_{r=0}^{n-1} A_r$$

$$\leqq B + M_2 n[\eta' A_n + K(\eta')A_0],$$

by (2.4.3). Hence

$$A_n[1 - nM_2\eta'] \leqq B + M_2 nK(\eta')A_0.$$

By suitable choice of η', η'' we can use (2.1.2) to deduce that

$$A_r \leqq \eta''(1 - nM_2\eta')^{-1}[B + nM_2K(\eta')A_0] + K(\eta'')A_0 \leqq \eta B + K(\eta)A_0 \qquad (r = 0, 1, \cdots, n).$$

(2.6) THEOREM 2. *Suppose* (1.1), (1.2), (1.3), (1.4) *are satisfied and* A_0, B *are finite. Then* $f_r(x)$ $(r = 0, 1, \cdots, n - 1)$ *converges to a finite limit* $f_r(\alpha)$, *respectively* $f_r(\alpha + a)$, *as* x *approaches* α, *respectively* $\alpha + a$, *and this limit is* 0 *if* α, *respectively* $\alpha + a$, *is not finite.*

By Theorem 1, $f_r(x)$ belongs to L^p for $r = 0, 1, \cdots, n$ and from (1.1) and Minkowski's inequality it follows that $f'_r(x)$ belongs to L^p for $r = 0, 1, \cdots, n - 1$. Now let (a_i, b_i) $(i = 1, 2, 3, \cdots)$ be any sequence of mutually exclusive subintervals of $(\alpha, \alpha + a)$ such that $\Re f_r(x)$ is non-vanishing (it is continuous) in every $a_i < x < b_i$. Then

$$\begin{aligned}
\sum_{i=1}^{\infty} \big|\, |\Re f_r(b_i)|^p - |\Re f_r(a_i)|^p \big| &\leqq \sum_{i=1}^{\infty} \int_{a_i}^{b_i} |(|\Re f_r(x)|^p)'|\, dx \\
&= \sum_{i=1}^{\infty} \int_{a_i}^{b_i} p\, |\Re f_r(x)|^{p-1}\, |\Re f'_r(x)|\, dx \\
&\leqq \int_{\alpha}^{\alpha+a} p\, |f_r(x)|^{p-1}\, |f'_{(r)}(x)|\, dx \\
&\leqq p \left[\int_{\alpha}^{\alpha+a} |f_r(x)|^p\, dx\right]^{\frac{p-1}{p}} \left[\int_{\alpha}^{\alpha+a} |f'_r(x)|^p\, dx\right]^{1/p},
\end{aligned}$$

by Hölder's inequality. Since $f_r(x)$ is continuous, it follows easily that $|\Re f_r(x)|^p$, and with it $|\Re f_r(x)|$, converge to finite limits as x approaches α, respectively $\alpha + a$, and that the limits are equal to 0 if α, respectively $\alpha + a$, is not finite. Since $\Re f_r(x)$ is real and continuous, the preceding statement holds for $\Re f_r(x)$ also. Similarly it holds for $\Im f_r(x)$. Hence it holds for $f_r(x)$.

3. The case $M_1 = 0$.

(3.1) If each $q_{r,j}(x)$, $q_r(x)$ vanishes in (1.1), the function $f_r(x)$ reduces to the r-th derivative of $f(x)$ and

$$A_r = \left[\int_{\alpha}^{\alpha+a} |f^r(x)|^p\, dx\right]^{1/p}. \tag{3.1.1}$$

In this case, we can prove more precise inequalities than those of Theorem 1. We shall prove the following results.

THEOREM 3. *Suppose that* $f(x)$ *is defined in* $(\alpha, \alpha + a)$ *and that its* $(n - 1)$*-th derivative function* $f^{n-1}(x)$ *exists and is absolutely continuous and that* A_0, A_n, *defined by* (3.1.1), *are finite. Then*

$$A_r \leqq K(n)\, [A_0 a^{-r} + A_0^{(n-r)/n} A_n^{r/n}] \qquad (r = 0, 1, \cdots, n), \tag{3.1.2}$$

where $K(n) < \infty$ *depends only on* n.

This has been proved by Hardy, Littlewood and Landau [1] in the case when a is infinite. The inequality for finite intervals enables us to deduce

THEOREM 4. *Suppose that A_r is defined for $r = 0, 1, \cdots, n$ by* (3.1.1), *that A_0 is finite, and that*

$$|p_r(x)| \leqq C^r \qquad (r = 1, 2, \cdots, n), \tag{3.1.3}$$

$$Tf(x) = f^n(x) + \sum_{r=0}^{n-1} p_{n-r}(x) f^r(x), \tag{3.1.4}$$

$$B = \left[\int_\alpha^{\alpha+a} |Tf(x)|^p \, dx\right]^{1/p} < \infty. \tag{3.1.5}$$

Then

$$A_r \leqq K(n) \left[A_0(C + a^{-1})^r + A_0^{(n-r)/n} B^{r/n}\right] \qquad (r = 0, 1, \cdots, n). \tag{3.1.6}$$

The finiteness of the A_r for the case $p = 2$ of Theorem 4 has been proved (by a somewhat different method) by J. von Neumann and the proof has been extended by one of the authors to apply to certain problems in the theory of differential operators [2]. The case $p = \infty$ (in which A_r is to be interpreted as the upper bound of $|f^r(x)|$ in $(\alpha, \alpha + a)$ was considered by Esclangon and Landau [3].[3]

We first prove two lemmas.

(3.2) LEMMA 3. *Let x be real, $a_r \geqq 0$, and*

$$x^n \leqq \sum_{r=0}^{n-1} a_{n-r} x^r.$$

Then

$$x \leqq \sum_{r=1}^{n} a_r^{1/r}.$$

For if

$$x > \sum_{r=1}^{n} a_r^{1/r},$$

we have also

$$x > a_r^{1/r} \qquad (r = 1, 2, \cdots, n),$$

so that

$$x^n > x^{n-1} \sum_{r=1}^{n} a_r^{1/r} = \sum_{r=1}^{n} x^{n-r} x^{r-1} a_r^{1/r}$$
$$> \sum_{r=1}^{n} x^{n-r} a_r = \sum_{r=0}^{n-1} a_{n-r} x^r.$$

(3.3) LEMMA 4. *With the hypotheses of Theorem* 3,

$$A_1 \leqq 24 A_0 a^{-1} + 2\sqrt{A_0 A_2}.$$

[3] See also R. P. Boas, Jr., this Journal, vol. 3(1937), pp. 637–646, especially footnote 1.

Putting $M_1 = 0$ in Lemma 2, we have

$$A_1 \leqq \left[\frac{2A_0}{h} + \frac{hA_2}{2}\right]\left[\frac{a-2h}{a-4h}\right] = \left[\frac{2A_0}{h} + \frac{hA_2}{2}\right]\left[1 + \frac{2h}{a-4h}\right].$$

Let $h = \min\left[\tfrac{1}{8}a, 2\sqrt{A_0/A_2}\right]$. Then either

$$h = 2\sqrt{A_0/A_2} \leqq \tfrac{1}{8}a,$$

$$\frac{2A_0}{h} + \frac{hA_2}{2} = 2\sqrt{A_0A_2}, \qquad \frac{2h}{a-4h} \leqq \frac{4h}{a} = \frac{8}{a}\sqrt{A_0/A_2};$$

or

$$h = \tfrac{1}{8}a \leqq 2\sqrt{A_0/A_2},$$

$$\frac{2A_0}{h} + \frac{hA_2}{2} \leqq \frac{16A_0}{a} + \sqrt{A_0A_2}, \qquad \frac{2h}{a-4h} = \frac{1}{2}.$$

In each case,

$$A_1 \leqq 24A_0a^{-1} + 2\sqrt{A_0A_2}.$$

(3.4) *Proof of Theorem* 3. The proof is by induction. We suppose that (3.1.2) holds with $n-1$ instead of n. That is,

$$A_r \leqq K[A_0a^{-r} + A_0^{1-r/(n-1)}A_{n-1}^{r/(n-1)}] \quad (r = 0, 1, \cdots, n-1). \tag{3.4.1}$$

In particular,

$$A_{n-2} \leqq K[A_0a^{-n+2} + A_0^{1/(n-1)}A_{n-1}^{(n-2)/(n-1)}]. \tag{3.4.2}$$

By Lemma 4,

$$A_{n-1} \leqq KA_{n-2}a^{-1} + KA_{n-2}^{1/2}A_n^{1/2}.$$

Hence, using (3.4.2),

$$A_{n-1} \leqq K[A_0a^{-n+1} + A_0^{1/(n-1)}A_{n-1}^{(n-2)/(n-1)}a^{-1}] \\ + KA_n^{1/2}[A_0^{1/2}a^{-(n-2)/2} + A_0^{1/[2(n-1)]}A_{n-1}^{(n-2)/[2(n-1)]}],$$

and an application of Lemma 3 gives[4]

$$\begin{aligned} A_{n-1} &\leqq K[A_0a^{-n+1} + A_n^{1/2}A_0^{1/2}a^{-(n-2)/2} + A_n^{(n-1)/n}A_0^{1/n}] \\ &\leqq K[A_0a^{-n+1} + A_0^{1/n}A_n^{(n-1)/n}], \end{aligned}$$

since[5]

$$\begin{aligned} A_n^{1/2}A_0^{1/2}a^{-(n-2)/2} &= (A_n^{(n-1)/n}A_0^{1/n})^{n/[2(n-1)]}(A_0a^{-n+1})^{(n-2)/[2(n-1)]} \\ &\leqq A_0a^{-n+1} + A_n^{(n-1)/n}A_0^{1/n}. \end{aligned}$$

[4] Choose $x = A_{n-1}^{1/[2(n-1)]}$ in Lemma 3 and use the inequality $(u + v + w)^r \leqq K(u^r + v^r + w^r)$ for $u, v, w \geqq 0$ and r integral.

[5] Use the inequality $u + v \geqq u^\alpha v^\beta$ for $\alpha, \beta, u, v \geqq 0$ and $\alpha + \beta = 1$.

This proves the inequality (3.1.2) for $r = n - 1$. (3.1.2) holds trivially if $r = n$, and for other values of r we can use (3.4.1) and deduce that

$$\begin{aligned} A_r &\leqq K[A_0 a^{-r} + A_0^{1-r/(n-1)} A_{n-1}^{r/(n-1)}] \\ &\leqq K[A_0 a^{-r} + A_0^{1-r/(n-1)}(A_0 a^{-n+1} + A_n^{(n-1)/n} A_0^{1/n})^{r/(n-1)}] \\ &\leqq K[A_0 a^{-r} + A_0^{1-r/n} A_n^{r/n}]. \end{aligned}$$

(3.5) *Proof of Theorem* 4. As in the proof of Theorem 1, we can suppose that the A_r's are finite. We have

$$|f^n(x)| \leqq |Tf(x)| + \sum_{r=0}^{n-1} C^{n-r} |f^r(x)|,$$

and using Minkowski's inequality and Theorem 3,

$$\begin{aligned} A_n &\leqq B + \sum_{r=0}^{n-1} C^{n-r} A_r \\ &\leqq B + K \sum_{r=0}^{n-1} C^{n-r}[A_0 a^{-r} + A_0^{(n-r)/n} A_n^{r/n}] \\ &\leqq B + KC^n A_0 \sum_{r=0}^{n-1} (aC)^{-r} + K \sum_{r=0}^{n-1} (CA_0^{1/n})^{n-r} A_n^{r/n} \\ &\leqq B + KA_0(C + a^{-1})^n + K \sum_{r=0}^{n-1} (CA_0^{1/n})^{n-r} A_n^{r/n}. \end{aligned}$$

It follows from Lemma 3 that

$$\begin{aligned} A_n^{1/n} &\leqq KCA_0^{1/n} + [B + KA_0(C + a^{-1})^n]^{1/n} \\ &\leqq KA_0^{1/n}\left(C + \frac{1}{a}\right) + KB^{1/n}. \end{aligned}$$

Using Theorem 3 again, we have

$$\begin{aligned} A_r &\leqq K\left[A_0 a^{-r} + A_0^{(n-r)/n} A_0^{r/n}\left(C + \frac{1}{a}\right)^r + A_0^{(n-r)/n} B^{r/n}\right] \\ &\leqq K[A_0(C + a^{-1})^r + A_0^{(n-r)/n} B^{r/n}]. \end{aligned}$$

4. Differential operators in L^p.

(4.1) Let $\mathfrak{D}$ be the linear space of those functions $f(x)$ belonging to $L^p(a, b)$ for which the relations (almost everywhere in the finite interval (a, b))

$$\begin{aligned} &f_0(x) = q_0(x)f(x), \\ &f_{r+1}(x) = f_r'(x) + q_{r+1}(x)f(x), \quad (r = 0, 1, \cdots, n-1), \end{aligned} \tag{4.1.1}$$

$$|q_0(x)| \geqq \epsilon > 0, \qquad |q_r(x)| \leqq C, \qquad (r = 0, 1, \cdots, n) \tag{4.1.2}$$

define functions $f_r(x)$ absolutely continuous for $r = 0, 1, \cdots, n-1$ and belonging to L^p for $r = n$. Let T be the quasi-differential operator in L^p with domain $\mathfrak{D}$,

$$Tf(x) = \sum_{r=0}^{n} p_r(x) f_r(x) + c(x)f(x) \tag{4.1.3}$$

with

$$|p_n(x)| \geqq \epsilon > 0, \quad |p_r(x)| \leqq C, \quad |c(x)| \leqq C \qquad (r = 0, 1, \cdots, n). \tag{4.1.4}$$

(Since $f(x)$ and $f_n(x)$ belong to L^p, (2.1.2) shows that $f_r(x)$ belongs to L^p for $r = 0, 1, \cdots, n$; the boundedness of $p_r(x)$, $c(x)$ then implies that $Tf(x)$ belongs to L^p for every $f(x)$ in $\mathfrak{D}$.)

Let $\mathfrak{D}_0$ be the set of the $f(x)$ in $\mathfrak{D}$ for which $f_r(a) = f_r(b) = 0$ for $r = 0, 1, \cdots, n-1$ (see Theorem 2). Let T_0 be the operator T restricted to the domain $\mathfrak{D}_0$.

We shall show that $\mathfrak{D}_0$ (and a fortiori $\mathfrak{D}$) is dense in L^p and we shall determine explicitly the adjoint and closure operators of T and T_0.

(4.2) Let $\mathfrak{D}^*$ be the linear space of the functions $g(x)$ belonging to $L^{p'}(a, b)$, where $\dfrac{1}{p} + \dfrac{1}{p'} = 1$, for which

$$\begin{aligned} g_0^*(x) &= q_0^*(x) g(x), \\ g_{r+1}^*(x) &= g_r^{*\prime}(x) + q_{r+1}^*(x) g(x) \qquad (r = 0, 1, \cdots, n-1) \end{aligned} \tag{4.2.1}$$

with

$$q_r^*(x) = (-1)^r \overline{p_{n-r}(x)} \qquad (r = 0, 1, \cdots, n) \tag{4.2.2}$$

define functions $g_r^*(x)$ absolutely continuous for $r = 0, 1, \cdots, n-1$ and belonging to $L^{p'}$ for $r = n$. Let T^* be the linear operator in $L^{p'}$ with domain $\mathfrak{D}^*$,

$$T^* g(x) = \sum_{r=0}^{n} p_r^*(x) g_r^*(x) + c^*(x) g(x), \tag{4.2.3}$$

where

$$c^*(x) = \overline{c(x)}, \qquad p_r^*(x) = (-1)^r \overline{q_{n-r}(x)}, \quad (r = 0, 1, \cdots, n). \tag{4.2.4}$$

Let $\mathfrak{D}_0^*$ be the set of $g(x)$ in $\mathfrak{D}^*$ for which $g_r^*(a) = g_r^*(b) = 0$ $(r = 0, 1, \cdots, n-1)$ and let T_0^* be the operator T^* restricted to the domain $\mathfrak{D}_0^*$.

THEOREM 5. *$\mathfrak{D}_0$ is dense in L^p and $\mathfrak{D}_0^*$ is dense in $L^{p'}$.*

THEOREM 6. *T_0^* and T^* are the adjoint operators of T and T_0, respectively. T and T_0 are closed operators.*

Theorems 5 and 6 have been proved for the case $p = 2$ by one of the authors [2]. These theorems show that the processes of taking closures and adjoint applied to the class of ordinary differential operators in L^p lead to the more general class of quasi-differential operators but that this latter class of operators is closed under these processes. A general discussion of linear operators in L^p is given by F. J. Murray [4].

We require the following lemmas.

(4.3) LEMMA 5. *If $f(x)$, $\varphi(x)$ are integrable over the finite interval (a, b) and $Q(x, z)$ is measurable and bounded on $a \leqq x, z \leqq b$, then*

$$f(x) = \int_a^x [\varphi(z) - Q(x, z)f(z)]\, dz \tag{4.3.1}$$

is equivalent to

$$f(x) = \int_a^x K(x, z)\varphi(z)\, dz \tag{4.3.2}$$

with

$$K(x, z) = K(Q, x, z) = 1 - \int_z^x Q(x, z_1)\, dz_1 + \int_z^x dz_1 \int_{z_1}^x Q(x, z_2)Q(z_2, z_1)\, dz_2$$

$$- \int_z^x dz_1 \int_{z_1}^x Q(z_2, z_1)\, dz_2 \int_{z_2}^x Q(x, z_3)Q(z_3, z_2)\, dz_3 + \cdots .$$

The series for $K(Q, x, z)$ can be obtained by successive substitutions for $f(z)$ in (4.3.1) with suitable changes in the order of integration. The uniform convergence of the infinite series and the verification of the lemma follow easily from the boundedness of $Q(x, z)$.

(4.4) LEMMA 6. *Let (a, b) be a finite interval and let $K_{j,r}(x, z)$ be defined by induction by*

$$\begin{aligned} &K_{0,1}(x, z) = K(Q, x, z) \quad \textit{with} \quad Q(x, z) = \frac{q_1(z)}{q_0(z)}, \\ &K_{r,r+1}(x, z) = K(Q, x, z) \quad \textit{with} \quad Q(x, z) = \int_z^x K_{0,r}(z_1, z)\frac{q_{r+1}(z_1)}{q_0(z_1)}\, dz_1 \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (r = 1, 2, \cdots, n-1), \\ &K_{j,r+1}(x, z) = \int_z^x K_{j,r}(x, z_1)K_{r,r+1}(z_1, z)\, dz_1 \\ &\qquad\qquad\qquad\qquad\qquad (j = 0, 1, \cdots, r-1;\, r = 1, 2, \cdots, n-1). \end{aligned} \tag{4.4.1}$$

The kernels $K_{j,r}$ have the following property. If $f(x)$ is in $\mathfrak{D}$ with $f_r(a) = 0$ $(r = 0, 1, \cdots, n-1)$, then

$$\begin{aligned} &f(x) = \frac{f_0(x)}{q_0(x)}, \\ &f_j(x) = \int_a^x K_{j,r}(x, z)f_r(z)\, dz \\ &\qquad\qquad (j = 0, 1, \cdots, r-1;\, r = 1, 2, \cdots, n). \end{aligned} \tag{4.4.2}$$

Conversely, if $\varphi(x)$ is any function belonging to L^p, then (4.4.2) with $f_n(x)$ replaced by $\varphi(x)$ determine a unique $f(x)$ in $\mathfrak{D}$ for which $f_n(x) = \varphi(x)$ and $f_r(a) = 0$ for $r = 0, 1, \cdots, n-1$.

The proof of this lemma follows from applying Lemma 5 to the relations

$$f_r(x) = \int_a^x [f_{r+1}(z) - q_{r+1}(z)f(z)]\, dz$$

and using suitable changes in the orders of integration.

LEMMA 7. *The $K_{j,r}(x, z)$ of Lemma 6 satisfy*

$$\frac{\partial^s K_{j,r}(x, z)}{\partial z^s} \text{ is bounded} \qquad (s = 0, 1, \cdots, r); \tag{4.4.3}$$

$$\left[\frac{\partial^s K_{j,r}(x, z)}{\partial z^s}\right]_{z=x} = \begin{cases} 0 & (s = 0, 1, \cdots, r-1;\ s \neq r-j-1), \\ \pm 1 & (s = r-j-1). \end{cases} \tag{4.4.4}$$

The proof is by induction. For $n = 1$ the lemma is easily verified. If now the lemma is proved for $1, 2, \cdots, n-1$, then it is easily proved for n by using (4.4.1) and induction on $j = n-1, n-2, \cdots, 3, 2, 1, 0$.

(4.5) LEMMA 8. *Let $\mathfrak{M}$ be any linear subspace of L^p and $\mathfrak{M}'$ the set of $\varphi(x)$ belonging to $L^{p'}$ for which*

$$\int_a^b \varphi(x)\overline{\psi(x)}\, dx = 0$$

for every $\psi(x)$ in $\mathfrak{M}$. Suppose $f(x)$ belongs to L^p and that

$$\int_a^b f(x)\overline{\varphi(x)}\, dx = 0$$

for every $\varphi(x)$ in $\mathfrak{M}'$. Then $f(x)$ is in the closure of $\mathfrak{M}$.

The proof of this theorem for complex L^p-space is due to Murray [4] (see Theorem 1.3 there). In particular, if $\mathfrak{M}$ consists of the single function $\psi(x) = 0$ (multiplied by an arbitrary numerical factor), then $f(x) = 0$ almost everywhere.

(4.6) LEMMA 9. *If $\varphi(x)$ belongs to $L^{p'}$ and*

$$\int_a^b f(x)\overline{\varphi(x)}\, dx = 0$$

for every $f(x)$ in $\mathfrak{D}_0$, then $\varphi(x) = 0$ almost everywhere.

By Lemma 6

$$\int_a^b \overline{\varphi(x)}\, dx \int_a^x K_{0,n}(x, z) f_n(z)\, dz = 0,$$

that is, by a suitable change in the order of integration,

$$\int_a^b f_n(z)\, dz \int_z^b K_{0,n}(x, z)\overline{\varphi(x)}\, dx = 0,$$

for every $f_n(z)$ in L^p for which

$$\int_a^b K_{r,n}(b, z) f_n(z)\, dz = 0 \qquad (r = 0, 1, \cdots, n-1).$$

Since the finite-dimensional linear subspace of functions of the form

$$\sum_{r=0}^{n-1} c_r K_{r,n}(b, z)$$

is closed, we have by Lemma 8,

$$\int_z^b K_{0,n}(x, z)\overline{\varphi(x)}\, dx = \sum_{r=0}^{n-1} c_r K_{r,n}(b, z).$$

Setting $z = b$ after partial differentiation r times with respect to z gives, with the help of Lemma 7, $c_r = 0$ $(r = 0, 1, \cdots, n-1)$. Differentiating n times with respect to z gives

$$\varphi(z) = \pm \int_z^b \frac{\overline{\partial^n K_{0,n}(x, z)}}{\partial z^n} \varphi(x)\, dx$$

a relation which implies, by Lemma 5 and (4.4.3), that $\varphi(x) = 0$ almost everywhere.

(4.7) LEMMA 10.

(4.7.1) $$\int_a^b Tf(x)\overline{g(x)}\, dx - \int_a^b f(x)\overline{T^*g(x)}\, dx = \sum_{r=0}^{n-1} (-1)^r [f_{n-r-1}(x) g_r^*(x)]_a^b$$

for all $f(x)$ in $\mathfrak{D}$ and $g(x)$ in $\mathfrak{D}^*$.

This identity, giving the extension of the well-known bilinear concomitant of ordinary differential operators[6] to quasi-differential operators, can be verified by repeated integration by parts.

(4.8) *Proof of Theorem* 5. From Lemmas 8 and 9 it follows that the closure of $\mathfrak{D}_0$ in L^p is L^p, that is, $\mathfrak{D}_0$ is dense in L^p. Similarly, $\mathfrak{D}_0^*$ is dense in $L^{p'}$.

(4.9) *Proof of Theorem* 6. Let T' be the adjoint operator to T. Then $T'g(x) = h(x)$ is equivalent to

$$g(x),\ h(x) \text{ are in } L^{p'}, \text{ and}$$

(4.9.1) $$\int_a^b Tf(x)\overline{g(x)}\, dx = \int_a^b f(x)\overline{h(x)}\, dx \quad \text{for all} \quad f(x) \text{ in } \mathfrak{D}.$$

By Lemma 6, (4.9.1) implies

$$\int_a^b \left\{ p_n(x)f_n(x) + \sum_{r=0}^{n-1} p_r(x) \int_a^x K_{r,n}(x, z)f_n(z)\, dz + \frac{c(x)}{q_0(x)} \int_a^x K_{0,n}(x, z)f_n(z)\, dz \right\} \overline{g(x)}\, dx$$
$$= \int_a^b \left\{ \frac{1}{q_0(x)} \int_a^x K_{0,n}(x, z)f_n(z)\, dz \right\} \overline{h(x)}\, dx$$

for arbitrary $f_n(x)$ in L^p. By suitable changes in the order of summations and integrations, which are easily justified, this can be written

$$\int_a^b f_n(x) \left\{ p_n(x)\overline{g(x)} + \sum_{r=0}^{n-1} \int_x^b K_{r,n}(z, x)p_r(z)\overline{g(z)}\, dz \right.$$
$$\left. + \int_x^b K_{0,n}(z, x) \frac{c(z)}{q_0(z)} \overline{g(z)}\, dz - \int_x^b K_{0,n}(z, x) \frac{\overline{h(z)}}{q_0(z)}\, dz \right\} = 0$$

[6] See E. L. Ince, *Ordinary Differential Equations*, pp. 123–124.

for arbitrary $f_n(x)$ in L^p. Using Lemma 8, we deduce that the expression in braces equals 0 almost everywhere. Repeated differentiation and the use of Lemma 7 show that $g(x)$ is in $\mathfrak{D}^*$ and $g_r^*(b) = 0$ $(r = 0, 1, \cdots, n-1)$. Interchanging the rôles of the end-points a, b we deduce that $g(x)$ is in $\mathfrak{D}_0^*$. Since $\mathfrak{D}_0$ is dense in L^p, T' is single-valued (see Murray, Theorem 2.6, [4]). Now Lemma 10 shows that $T'g(x) = T_0^*g(x)$. From the same Lemma 10 it is plain that $T_0^*g(x) = h(x)$ implies that $T'g(x)$ is defined and equals $h(x)$. Hence T_0^* is the adjoint operator to T.

Applying the preceding paragraph to T^* (in place of T), and using the theorem that the adjoint of the adjoint is the closure (Murray, Theorem 2.6, [4]), we obtain that T^* is the adjoint operator to T_0 and that T and T_0 are closed operators.

(4.10) The extension of Theorems 5 and 6 to infinite intervals and the discussion of linear boundary conditions can be carried out here in precisely in the same way as has been done in the case $p = 2$ in [2].

References

1. G. H. Hardy, E. Landau, and J. E. Littlewood, *Some inequalities satisfied by the integrals or derivatives of analytic functions*, Math. Zeitschrift, vol. 39(1935), pp. 677–695.

2. I. Halperin, *Closures and adjoints of linear differential operators*, Annals of Math., vol. 38(1937), pp. 880–919.

3. E. Landau, *Über einen Satz von Herrn Esclangon*, Math. Annalen, vol. 102(1929), pp. 177–188.

4. F. J. Murray, *Linear transformations in L^p, $p > 1$*, Trans. Amer. Math. Soc., vol. 39 (1936), pp. 83–100.

Harvard University.

INTERIOR SURFACE TRANSFORMATIONS

By G. T. Whyburn

In a previous paper[1] it was shown that when a 2-dimensional manifold A (with or without boundary curves) undergoes a light interior transformation $T(A) = B$, the resulting image B is likewise a 2-dimensional manifold. In this paper it will be shown that under these circumstances the Euler characteristics of the original and resulting manifolds are connected by a simple numerical relationship involving integers dependent only on A, B and the transformation T (see §2). Numerous examples and applications of this result follow in §§3 and 4. In the concluding section there is developed a method which effects the extension of this as well as other results concerning interior light transformations to the case of 2-dimensional pseudo-manifolds.

1. We consider a light interior transformation $T(A) = B$, where A (hence also[1] B) is a compact 2-dimensional manifold. Let α and β denote the boundaries (if any) of A and B, respectively. By a previous theorem[2] it follows that there exists an integer k such that the inverse of every point in B consists of k or fewer points. We define the least such integer k to be the *degree* of T. In other words, the degree k of T is the maximum multiplicity of T. Also from the theorem just cited it follows that there is only a finite number of points of $B - \beta$ whose inverse contains a point where T is not locally topological. Thus if Z denotes the set of all such points of $B - \beta$, it follows that on the set $A - T^{-1}(Z) - T^{-1}(\beta)$, T is locally topological; and since this set is connected, T must[3] be exactly k to 1 on this set. Thus k might also be defined as the multiplicity of T on the set $A - T^{-1}(\beta + Z)$.

Throughout this section the letters used above will retain their significance as there defined. Also $\chi(X)$ will stand for the Euler characteristic of the complex or surface X. We proceed to establish two lemmas.

Lemma 1. *If $N \subset B$ is a graph dividing B into only a finite number of components, then $T^{-1}(N)$ is a graph.*

Proof. Since all but a finite number of points of N are of order 2, $T^{-1}(N)$ has at most a finite number of end points. Since $A - T^{-1}(N)$ has only a finite number of components (each of which maps onto a component of $B - N$), it follows that $T^{-1}(N)$ contains only a finite number of simple closed curves.

Received July 9, 1938.

[1] See my paper in the American Journal of Mathematics, vol. 60(1938), pp. 477–490.

[2] Ibid., Theorem (5.2).

[3] See Eilenberg, Fundamenta Mathematicae, vol. 24(1935), p. 36.

Thus each cyclic element of a component C of $T^{-1}(N)$ is a graph, and as C can have only a finite number of nodes, C is a graph and so is $T^{-1}(N)$.

In particular, $T^{-1}(\beta)$ is a graph, since $B - \beta$ is connected.

LEMMA 2. *There exist simplicial subdivisions K and H of A and B, respectively, such that each simplex of K maps topologically onto a simplex of H.*

Proof. Let Z be the set of all points of $B - \beta$ whose inverse contains a point at which T is not locally topological. For each point $q \in T^{-1}(Z)$, where T is not locally topological, let (see footnote 2) E_q be a 2-cell neighborhood of q on which T is equivalent to the transformation $w = z^k$ on $|z| \leqq 1$, and let G_q be the graph consisting of the part in E_q of the inverse of the set corresponding to $v = 0$, $-1 \leqq u \leqq 1$, where $w = u + iv$. As the number of points in $T^{-1}(Z)$ is finite, we may suppose the graphs G_q are disjoint and do not intersect $T^{-1}(\beta)$. Let G be the graph consisting of $T^{-1}(\beta)$, $T^{-1}(Z)$ and all the graphs G_q. Let $W = T^{-1}(\beta) + T^{-1}(Z)$.

(i) There exists a $d > 0$ such that if E is any region in A not intersecting G and such that $\rho(E, W) < d > \delta(E)$, then T is topological on $\bar{E}$.

Proof of (i). Let $Q = q_1 + q_2 + \cdots + q_m$ be the points of $T^{-1}(Z)$ where T is not locally topological. Now (see footnote 2) the set $V = T^{-1}(Z) - Q + T^{-1}(\beta)$ can be covered by a finite number of 2-cells $D_1 + D_2 + \cdots + D_p$ such that the part, if any, of $T^{-1}(\beta)$ within D_i divides D_i into a finite number of 2-cells on each of which T is topological. Now there exists a $d_0 > 0$ such that any subset A' of A such that $\delta(A') < d_0$ and $\rho(V, A') < d_0$ lies wholly in some one of the sets D_i. For each $j \leqq m$, let d_j be less than half the distance from q_j to the boundary of E_{q_j}.

Finally, let $d = \min d_i$, $0 \leqq i \leqq m$. If E is any region in A of diameter $< d$ not intersecting G and such that $\rho(E, W) < d$, it is seen at once that E lies wholly in some one of the sets D_i or in one of the sets E_{q_j}; and, in either case, since the interior of E does not intersect G, it follows that T is topological on $\bar{E}$.

(ii) There exists a number $e > 0$ such that if K is any subdivision (simplicial or cellular) of A of norm $< e$ whose 1-dimensional structure includes all points of the graph G, then on any simplex of K, T is topological.

For if we take a sufficiently small neighborhood U of $\beta + Z$ in B, we will have $W \subset T^{-1}(U) \subset V_d(W)$. Hence on the set $T^{-1}(B - U)$, T is locally topological. Accordingly, there exists a number $f > 0$ such that on any subset of $T^{-1}(B - U)$ of diameter $< f$, T is topological. Let $e = \min(d, f)$. Then if K is any subdivision of A of norm $< e$ whose 1-dimensional structure includes G and if E is any 2-cell of K, we have either $E \subset T^{-1}(B - U)$ or $\rho(E, W) < d$; and since the interior of E cannot intersect G, it follows in either case that T is topological on E. Hence surely T is topological on any cell of K.

Since T is light, we have

(iii) There exists a number $h > 0$ such that if H' is any graph in B including $\beta + Z$ and such that every component of $B - H'$ is of diameter $< h$, the diameter of every component of $T^{-1}(B - H') = A - T^{-1}(H')$ is $< e$.

Let H be a simplicial subdivision of B whose 1-dimensional structure H' is such a graph.

Since every point of Z is a vertex of H, it follows that $T^{-1}(H') = K'$ contains a graph G_q about each point q_i of q. Thus by (iii) and (ii) the closure of each component of $A - K'$ is a 2-cell mapping topologically onto a simplex F of H. Thus if for each such E we select edges and vertices by the (topological) transformation $T^{-1}(F) = E$, it is clear that we obtain a simplicial subdivision K of A with 1-dimensional structure K' such that T is topological on each simplex of K.

COROLLARY. *If B is a closed orientable surface, so also is A.*[4]

For we have only to assign to each 2-simplex in K the orientation of its image simplex in H, where H is "coherently" oriented.[5] Then since each 1-simplex in K is on exactly two (see reference in footnote 1) 2-simplexes and these have distinct images, it follows that K is coherently oriented.

Thus non-orientability is invariant under light interior transformations between closed surfaces. That this is not true for surfaces with boundary will be seen below in §3, example (iii).

2. THEOREM. *If A is a 2-dimensional manifold and $T(A) = B$ is a light interior transformation of degree k, then*

$$k\chi(B) - \chi(A) = kr - n - m,$$

where r and n are the numbers of points in Y and $T^{-1}(Y)$, respectively, where Y is the set of all $y \in B$ such that $T^{-1}(y)$ contains either a branch point of $T^{-1}(\beta)$ or a point of $A - T^{-1}(\beta)$ at which T is not locally topological, and where m is the number of components of $T^{-1}(\beta) - T^{-1}(Y)\cdot T^{-1}(\beta) - \alpha$ which are not simple closed curves.

Proof. Let K and H be subdivisions of A and B given by Lemma 2. Let α^i and β^i ($i = 0, 1, 2$) be the number of i-dimensional simplexes in K and H, respectively.

Since T is k to 1 on the set $A - T^{-1}(Z) - T^{-1}(\beta)$, we have

(i) $$\alpha^2 = k\beta^2.$$

Since for any 1-simplex x^1 of K which is in $T^{-1}(\beta)$ but not in α the two 2-simplexes on x^1 have the same image, it follows that if t is the total number of such 1-simplexes x^1 in K, we have

(ii) $$\alpha^1 = k\beta^1 - t.$$

For the same reason, if s is the number of vertices of K on the set $T^{-1}(\beta) - T^{-1}(Y)\cdot T^{-1}(\beta) - \alpha$, we have

(iii) $$\alpha^0 = k\beta^0 - (kr - n) - s.$$

Now let $E_1, E_2, \cdots, E_m, E_{m+1}, \cdots, E_q$ be the components of $T^{-1}(\beta) - T^{-1}(Y)\cdot T^{-1}(\beta) - \alpha$, where $E_1, E_2, \cdots, E_m$ are open arcs and $E_{m+1}, E_{m+2}, \cdots, E_q$ are simple closed curves.

[4] Compare with Stoïlow, Compositio Mathematica, vol. 3(1936), pp. 435–440.

[5] For a definition of this term, see Alexandroff-Hopf, *Topologie*, Berlin, 1935.

Let t_i and s_i be the number of edges and vertices respectively of K on E_i $(0 \leqq i \leqq q)$. Then

$$\sum_1^q t_i = t, \qquad \sum_1^q s_i = s.$$

Further, for $i > m$, $t_i = s_i$; and for $i \leqq m$, $t_i = s_i + 1$, whence

$$\text{(iv)} \qquad \begin{aligned} t - s = \sum_1^q (t_i - s_i) &= \sum_1^m (t_i - s_i) + \sum_{m+1}^q (t_i - s_i) \\ &= \sum_1^m 1 + 0 = m. \end{aligned}$$

Now, by (i)–(iii), we have

$$\alpha^2 - \alpha^1 + \alpha^0 = k(\beta^2 - \beta^1 + \beta^0) - (kr - n) + t - s.$$

Whence, by (iv),

$$\chi(A) = k\chi(B) - (kr - n) + m,$$

or

$$k\chi(B) - \chi(A) = kr - n - m.$$

3. **Examples.** (i) Let A be a sphere with center at the origin in 3-space with a cylindrical coördinate system (ρ, θ, z). Let T be the transformation of A into a sphere B effected by identifying the points (ρ, θ, z) and $(\rho, \theta + 2\pi/k, z)$, where k is a positive integer. We have $k = k$, $r = n = 2$, $m = 0$, so that both sides of the equation in our theorem reduce to $2k - 2$.

(ii) Let T_1 be the transformation of a torus A into a circular ring R effected by merely flattening the torus onto a plane which cuts it into two rings. Let T_2 be the transformation of R into a 2-cell B, effected by folding R across a line cutting R into two 2-cells. Finally, let $T = T_2T_1$. For T_1 we have: $k = 2$, $r = n = m = \chi(A) = \chi(R) = 0$. For T_2: $k = 2$, $r = n = 4$, $m = 2$, $\chi(R) = 0$, $\chi(B) = 1$. For T: $k = 4$, $r = n = 4$, $m = 8$, $\chi(A) = 0$, $\chi(B) = 1$.

(iii) Let T be the transformation of a projective plane A into a 2-cell B, effected as follows. Cut A into a Möbius band M and a 2-cell E, each bounded by a simple closed curve J. Regard M as generated by an interval ab moving "parallel" to itself with its ends on J. Now in M identify points on ab equidistant from the midpoint of ab; and on E let T be equivalent to the transformation $w = z^2$ on $|z| \leqq 1$. We have: $k = 2$, $r = n = 1$, $m = 0$, $\chi(A) = \chi(B) = 1$.

(iv) Let T be a transformation of a torus A into a sphere B effected as follows. Cut A into a 2-cell E and a "handle" H each bounded by a simple closed curve J. On E let T be equivalent to $w = z^2$ on $|z| \leqq 1$. Let H be mapped into a 2-cell F as follows. On B select three points u', v', w' inside F and join each to the edge of F by 2 intervals, thus dividing F into 4 regions. On H select three points u, v, w which will be branch points of order 2; for convenience

these may be taken with u and v opposite points on one generator of the torus and w opposite v on the other generator. Now join each of these points on H to J by 4 intervals so that H is divided into 8 regions. If selected properly, T can be defined so as to map these in pairs topologically onto the four regions of F. For this transformation we have: $k = 2$, $r = n = 4$, $m = 0$, $\chi(A) = 0$, $\chi(B) = 2$.

4. **Applications.**

(i) *If* $m = 0$ *[as for example when* $\alpha = T^{-1}(\beta)$*] and* $\chi(A) = \chi(B)$, *either* T *is topological or* $r \geqq \chi(A)$.

For in this case our equation in §2 takes the form

$$(k - 1)(\chi(A) - r) = r - n;$$

and since the right member is $\leqq 0$, we have either $k = 1$ or $r \geqq \chi(A)$.

Thus, for example, any non-topological light interior transformation of a sphere into a sphere has at least two singular points in the image sphere [see example (i)]; and any such transformation of a projective plane into a 2-cell or another projective plane or of a 2-cell into a 2-cell [see example (iii)] has at least 1 singular point in the image set.

(ii) *If* $\chi(A) = \chi(B) < 0$ *and* $m = 0$, *the transformation is necessarily topological.*

For our equation takes the form

$$(k - 1)\chi(A) = kr - n;$$

and as the right member is $\geqq 0$ and $\chi(A) < 0$, it follows that $k = 1$. Obviously this conclusion does not hold for surfaces of characteristic $\geqq 0$.

(iii) *If* $kr = n$ *and* $m = 0$ *(as for example when* T *is locally topological*[6]*),*

$$k\chi(B) = \chi(A).$$

Thus $\chi(A)$ *and* $\chi(B)$ *vanish or fail to vanish together and if* $\chi(A) = \chi(B) \neq 0$, T *is necessarily topological.*

It is interesting to note that the conditions imposed here do not restrict T to a local homeomorphism—see the transformation T_1 under example (ii) in §3.

For the surfaces of positive characteristic we have the following striking consequences. If A is a sphere, $\chi(A) = 2$ and hence $k = 1$ or 2 and $\chi(B) =$ 1 or 2. Thus B is either a sphere, projective plane or 2-cell and all are possible (a known result, see my paper in footnote 1) under such a transformation. If B is a sphere, $k = 1$ and the transformation is topological; if B is a projective plane, then $k = 2$ and T is locally topological and exactly (2, 1); if B is a 2-cell, $k = 2$ and T is not locally topological. Thus *the only exactly* $(k, 1)$ *interior transformations on a sphere are the topological transformation into another sphere*

[6] This conclusion was established in the case of a local homeomorphism on a linear graph by A. D. Wallace. See his abstract in the Bulletin of the American Mathematical Society, vol. 44(1938), p. 202.

and the (2, 1) *non-singular transformation into a projective plane* (identifying diametrically opposite points). For no value of k other than 1 and 2 does there exist a $(k, 1)$ interior transformation on a sphere.

If A is a projective plane or a 2-cell, $\chi(A) = 1$ and hence $\chi(B) = k = 1$. Thus the only transformation satisfying our conditions on A is a homeomorphism. In particular, it follows that *there exists no exactly* $(k, 1)$ *interior transformation on a projective plane or a 2-cell other than a homeomorphism.*

For surfaces of characteristic $\leqq 0$, of course, no such conclusions hold. For example, it is easy to define transformations of a torus into a torus or of a circular ring into a circular ring which are locally topological and exactly $(k, 1)$ for every positive integral value of k. Also, it is easily seen that for any $n = 0, 1, 2, \cdots$ and any $k = 1, 2, \cdots$ we can map the closed orientable surface of characteristic $-2kn$ into the one of characteristic $-2n$ by a $(k, 1)$ local homeomorphism.

In the case of a 2-cell or the projective plane, even without the assumption that $kr = n$, we have $k - 1 = kr - n$; thus $r = 1$ gives $n = 1$ so that if there is only 1 branch point in B, the same holds for A; and in the case of a 2-cell, T is equivalent to $w = z^k$ on $|z| \leqq 1$.

5. **Extension to pseudo-manifolds.** In this section it will be shown that any light interior transformation on a 2-dimensional pseudo-manifold (see footnote 5) reduces essentially to a light interior transformation defined on a manifold. This reduction will be made with the aid of a transformation previously studied by the author[7] which makes use of the relative distance space introduced by Mazurkiewicz.[8] If M is a connected metric space of finite diameter, the relative distance space M^* of M consists of the same points as M, but the metric $\rho^*(x, y)$ is defined, for $x,y \in M^*$, as the greatest lower bound of $[\delta(C)]$, where C is any connected subset of M containing $x + y$.

In the author's paper just cited it was shown that if M has property S (i.e., M is the sum of a finite number of arbitrarily small connected sets), and if, in general, for any metric space X, X_c denotes the space obtained by "completing" X, then the spaces M_c and M_c^* are compact and the change of metric from M^* to M generates a continuous transformation

$$W(M_c^*) = M_c$$

which maps M^* onto M topologically and uniformly continuously and maps $M_c^* - M$ onto $M_c - M$.

Lemma. *Let R be connected and have property S and let $W(R_c^*) = R_c$, where R^* is the relative distance space for R. If for any $p \in R_c - R$ and any $\epsilon > 0$ we let $n(\epsilon, p)$ be the number of components of $R \cdot V_{\epsilon/3}(p)$ having p for a limit point, then $W^{-1}(p)$ is the sum of $n(\epsilon, p)$ closed sets each of diameter $< \epsilon$ and any two of*

[7] See my paper in the American Journal of Mathematics, vol. 54(1932), pp. 367–376.

[8] See Fundamenta Mathematicae, vol. 1(1920), pp. 167–168.

which are at a distance apart $> \frac{1}{4}\epsilon$. *Thus* W *is light and if* $\lim_{\epsilon \to 0} n(\epsilon, p) = n(p)$, *then* $n(p)$ *is equal to the number of points in* $W^{-1}(p)$ *if either of these numbers is finite.*

Proof. Let $R_1, R_2, \cdots, R_{n(\epsilon,p)}$ be the components of $R \cdot V_{\epsilon/3}(p)$ having p as a limit point. For each i, let F_i be the set of all points x of $W^{-1}(p)$ such that if $x_1, x_2, \cdots$ is any sequence in R^* converging to x, then almost all of the points $W(x_1), W(x_2), \cdots$ are in R_i and of course the sequence converges to p. Since for any $x \in W^{-1}(p)$ and any sequence $x_k \to x$, where $x_k \in R^*$ for all k, there exists an i such that almost all $W(x_k)$ lie in R_i (because we have $\rho^*(y_i, y_j \geqq \frac{1}{6}\epsilon$ for $y_i \in R_i \cdot V_{\epsilon/6}(p)$, $y_j \in R_j \cdot V_{\epsilon/6}(p))$, it follows that

$$W^{-1}(p) = F_1 + F_2 + \cdots + F_{n(\epsilon,p)}.$$

Now if $x \in F_i$, $y \in F_j$, we have

$$\rho^*(x, y) = \lim \rho^*(x_k, y_k) = \lim \rho^*[W(x_k), W(y_k)] \geqq \tfrac{1}{4}\epsilon,$$

where $x_k \to x$, $y_k \to y$, $x_k, y_k \in R^*$. Hence

$$\rho(F_i, F_j) \geqq \tfrac{1}{4}\epsilon.$$

Thus since $W^{-1}(p)$ is closed, each F_i is closed.

Theorem. *Let* $T(A) = B$ *be interior and light, where* A *is a compact 2-dimensional pseudo-manifold. There exist compact 2-dimensional manifolds* A' *and* B' *and continuous transformations*

$$W(A') = A, \; T'(A') = B', \; Z(B') = B$$

such that

(1) W *is topological except at* f *points which map into* s *points of* A,
(2) Z *is topological except at* e *points which map into* t *points of* B,
(3) T' *is interior and light, and*
(4) $ZT'W^{-1} \equiv T$.

Proof. Let P denote the (finite) set of all local separating points of A and let $R = A - P$. Then R is connected and has property S. Let R^* be the relative distance space for R and set $R_c^* = A'$. Then since $R_c = A$, we must have

$$W(A') = A.$$

Furthermore, since for each $p \in P$ there exists a closed neighborhood of p which is the sum of a finite number n_p of 2-cells intersecting by pairs only in p, it follows by the preceding lemma that there are exactly n_p points in $W^{-1}(p)$ and each of these points has a closed 2-cell neighborhood. Thus since W is topological on $A' - W^{-1}(P)$, it follows that A' is a 2-dimensional manifold and assertion (1) holds.

The transformation $T'(A') = B'$ is given by the following continuous decomposition G of A' into disjoint closed sets G_x. For any $x \in A'$ such that

$W(x)$ is not a local separating point of A, let

$$G_x = W^{-1}T^{-1}T(x).$$

For an $x \in A'$ such that $W(x)$ is a local separating point of A, let E_x be a 2-cell in A' containing x so small that W is topological on E_x. Let $x_1, x_2, \cdots$ be a sequence of points in E converging to x, and let $G_x = \lim (G_{x_i})$. Clearly G_x is independent of the sequence x_i, since it consists merely of all $y \in A'$ such that

$$TW(y) = TW(x) = z$$

and $TW(E_y)$ is essentially the same as $TW(E_x)$, i.e., z is interior to the common part of these sets relative to their sum. For the same reason it follows that any two distinct sets G_x are disjoint and that the decomposition G of A' into the sets G_x is continuous. Let B' be the hyperspace of this decomposition and let

$$T'(A') = B'$$

be the associated (light interior) transformation. Then B' is a 2-dimensional manifold, since (see footnote 1) T' is interior and light.

Finally, if for each $x \in B'$ we define

$$Z(x) = TWT'^{-1}(x),$$

clearly we get $Z(B') = B$ and (4) holds. Also, from the definitions of Z and T' it follows that if $x,y \in B' - T'[W^{-1}(P)]$, $x \neq y$, we have

$$Z(x) \neq Z(y).$$

Hence (2) holds and our theorem is proved.

COROLLARY. *The image of a 2-dimensional pseudo-manifold under any light interior transformation is itself a 2-dimensional pseudo-manifold.*

Now to extend our theorem of §2 to pseudo-manifolds, we let $T(A) = B$ be a light interior transformation where A (hence also B) is a 2-dimensional pseudo-manifold. Let $\alpha, \beta, k, Y, r, n, m$ be defined for T exactly as in §§1 and 2, where now the s and t points of A and B mentioned in the preceding theorem are included in the n and r points of A and B, respectively, given by $T^{-1}(Y)$ and Y. (Note that the existence of the integers k, r, n, m follows from their existence in the manifold case by virtue of the preceding theorem.)

Now from the transformations W and Z we have by virtue of (1) and (2)

$$\text{(i)}\qquad \begin{aligned} \chi(A') &= \chi(A) + f - s, \\ \chi(B') &= \chi(B) + e - t. \end{aligned}$$

By §2, the transformation $T'(A') = B'$ gives

$$\text{(ii)}\qquad \begin{aligned} k\chi(B') - \chi(A') &= k(r - t + e) - (n - s + f) - m \\ &= k(r - t) - (n - s) + ke - f - m. \end{aligned}$$

Now, by (i),

$$k\chi(B') - \chi(A') = k\chi(B) - \chi(A) + ke - f - kt + s,$$

whence, by (ii),

$$\begin{aligned} k\chi(B) - \chi(A) &= k(r - t) - (n - s) - m + kt - s \\ &= kr - n - m. \end{aligned}$$

This is identically the same relation as in the theorem of §2.

Thus we have the

THEOREM. *If $T(A) = B$ is interior and light, where A is a 2-dimensional pseudo-manifold, then B is a 2-dimensional pseudo-manifold, and if k, r, n, m are defined as in* §§1, 2, *we have*

$$k\chi(B) - \chi(A) = kr - n - m.$$

UNIVERSITY OF VIRGINIA.

FUNCTIONS OF INTEGRABLE SQUARE IN SEVERAL COMPLEX VARIABLES

By S. Bochner

As in two previous notes[1] we consider in the space C_k of k complex variables

$$z = (z_1, \cdots, z_k), \qquad z_\kappa = x_\kappa + iy_\kappa,$$

point sets of a special nature which we call *tubes*. A point set T of C_k is a tube, if there exists a point set S in the space R_k of real variables $x = (x_1, \cdots, x_k)$ such that T consists of all k-dimensional planes

$$x_\kappa = x_\kappa^0 \qquad (-\infty < y_\kappa < \infty;\ \kappa = 1, \cdots, k) \tag{1}$$

for which $(x_1^0, \cdots, x_k^0)$ is any point of S. The set S is called the basis of T, and we also denote T more explicitly by T_S. The tube T_S is open or closed in C_k if and only if S is open or closed in R_k; it is convex if and only if S is convex, and the convex hull[2] $\tilde{T}$ of a tube T is again a tube whose basis $\tilde{S}$ is the convex hull of S.

We say that a function $f(z) = f(z_1, \cdots, z_k)$ is of integrable square in T if the function

$$f_x(y) = f(x_1 + iy_1, \cdots, x_k + iy_k)$$

belongs to the Lebesgue class L_2 over the y-space, for every $x \subset S$, and if moreover there exists a constant K such that

$$\int_{-\infty}^{\infty}\cdots\int |f_x(y)|^2\, dv_y \leqq K, \tag{2}$$

for all $x \subset S$, the symbol dv_y denoting the Euclidean volume element $dy_1 \cdots dy_k$.

In our first note we proved the following theorem. If $f(z)$ is analytic and of integrable square in an open tube T, then it also exists and is of integrable square in $\tilde{T}$. In the present paper we shall extend this theorem to the case of tubes which are not necessarily open.

Assumptions. (1) *The basis S is such that any two points P, Q of S have a finite Euclidean distance $D(P, Q)$ on S in the following sense. Corresponding to*

Received May 6, 1938.

[1] S. Bochner, *Bounded analytic functions in several variables and multiple Laplace integrals*, American Journal of Math., vol. 59(1937), pp. 731–738; *A theorem on analytic continuation of functions in several variables*, Annals of Math., vol. 39(1938), pp. 14–19. I am indebted to H. Behnke for pointing out to me that the theorem of the second note can be proved in a much simpler fashion. See K. Stein, *Zur Theorie der Funktionen mehrerer komplexer Veränderlichen*, Math. Annalen, vol. 114(1937), p. 557.

[2] This is the smallest convex set containing T; it is not necessarily closed.

any $\epsilon > 0$ there exist points $P_0, P_1, \cdots, P_n$ in S, with the following properties: (α) $P_0 = P$, $P_n = Q$; (β) the Euclidean distance $\overline{P_\nu P_{\nu+1}}$ is $< \epsilon$, for $\nu = 0, 1, \cdots, n-1$; and

$$(\gamma) \qquad \overline{P_0P_1} + \overline{P_1P_2} + \cdots + \overline{P_{n-1}P_n} \leqq D(P, Q).$$

(2) *The convex hull $\tilde{S}$ of S is k-dimensional.*

(3) *The function $f(z)$ is analytic and bounded in some ($2k$-dimensional, open) domain*[3] *U of C_k containing T_S, and is of integrable square in T_S.*

Assertion. *The function $f(z)$ exists and is of integrable square in the convex hull $\tilde{T}$ of T. Also, it satisfies relation (2) in $\tilde{S}$ for the same constant K as in S.*

This theorem includes our previous theorem bearing on open tubes. In fact, if T is open, assumptions (1) and (2) are trivially fulfilled, and assumption (3) is satisfied for $U = T$, subject to the inessential qualification that $f(z)$ is not necessarily bounded in the whole tube T but is so in every closed tube interior to it.

For the proof of the theorem we may assume that S is bounded. Otherwise we take a fixed point O of S, a sequence of spheres Σ_n in C_k with centers at O whose radii tend to infinity, and, for each n, the set S_n consisting of those points of S which, in the sense of assumption (1), can be connected with the point O within Σ_n. Obviously S_n tends to S, and T_{S_n} tends to T_S. Thus it is sufficient to prove our theorem for $S = S_n$.

Another specialization is less trivial, and will be justified afterwards; we assume that

$$(3) \quad f(z_1, \cdots, z_k) = O((y_1^2 + \cdots + y_k^2)^{-\frac{1}{2}(k+1)}) \qquad \text{as } y_1^2 + \cdots + y_k^2 \to \infty,$$

uniformly for $z \subset U$.

As a consequence of (3), there exists a constant A such that

$$(4) \qquad f(z_1, \cdots, x_\kappa \pm ia, \cdots, z_k) \leqq Aa^{-k-1}$$

for $\kappa = 1, \cdots, k$, and

$$(z_1, \cdots, x_\kappa \pm ia, \cdots, z_k) \subset U.$$

Corresponding to any $a > 0$ there exists in R_k an (open) neighborhood S_a of S such that the point set

$$(x_1, \cdots, x_k) \subset S_a \qquad (-a < y_\kappa < a; \kappa = 1, \cdots, k)$$

of C_k is contained in U. We may assume that S_a is contained in a sphere

$$(5) \qquad x_1^2 + \cdots + x_k^2 \leqq \alpha^2$$

whose radius α is independent of a. We now introduce the function

$$\varphi_a(x, t) = \int_{-a}^{a} \cdots \int_{-a}^{a} f_x(y) \cdot \exp\left[\sum_{\kappa=1}^{k} (x_\kappa + iy_\kappa)t_\kappa\right] dv_y$$

[3] Which need not contain an open tube.

for $x \subset S_a$ and arbitrary real points $t = (t_1, \cdots, t_k)$. We are interested in the partial derivative of this function with respect to x_1. The integrand on the right is an analytic function of $x_1 + iy_1$, hence its derivative with respect to x_1 is, but for the factor $-i$, its partial derivative with respect to y_1, thus

$$\frac{\partial \varphi_a(x, t)}{\partial x_1} = -i \int_{-a}^{a} \cdots \int_{-a}^{a} \frac{\partial}{\partial y_1} \left\{ f_x(y) \exp \left[\sum_{\kappa=1}^{k} (x_\kappa + iy_\kappa)t_\kappa \right] \right\} dv_y .$$

Carrying out the integration with respect to y_1, we obtain

$$-i \int_{-a}^{a} \cdots \int_{-a}^{a} f_x(a, y_2, \cdots, y_k) \exp \left[(x_1 + ia)t_1 + \sum_{\kappa=2}^{k} (x_\kappa + iy_\kappa)t_\kappa \right] dy_2 \cdots dy_k$$

and a similar term with $-a$ instead of a. Using (4), we obtain, for $x \subset S_a$,

$$\left| \frac{\partial \varphi_a(x, t)}{\partial x_1} \right| \leqq A(\alpha, t)a^{-1}$$

and in general,

$$\left| \frac{\partial \varphi_a(x, t)}{\partial x_\kappa} \right| \leqq A(\alpha, t)a^{-1} \qquad (\kappa = 1, \cdots, k),$$

where $A(\alpha, t)$ is independent of a. If $x' = (x_1', \cdots, x_k')$ and $x'' = (x_1'', \cdots, x_k'')$ are any two points of S, and if the connecting segment

$$x_\kappa = \rho x_\kappa' + (1 - \rho)x_\kappa'' \qquad (0 \leqq \rho \leqq 1)$$

lies in S_a, then

$$\begin{aligned} | \varphi_a(x', t) - \varphi_a(x'', t) | &\leqq \sum_{\kappa=1}^{k} \max_{0 \leqq \rho \leqq 1} \left| \frac{\partial \varphi_a(x, t)}{\partial x_\kappa} \right| \cdot | x_\kappa' - x_\kappa'' | \\ &\leqq kA(\alpha, t) \left[\sum_{\kappa=1}^{k} (x' - x'')^2 \right]^{\frac{1}{2}} \cdot a^{-1}. \end{aligned}$$

Consequently, by assumption (1),

$$| \varphi_a(x', t) - \varphi_a(x'', t) | \leqq k \cdot A(\alpha, t) D(x', x'') \cdot a^{-1} ,$$

for any two points x', x'' of S. Letting $a \to \infty$, we conclude that the integral

$$\int_{-\infty}^{\infty} \!\! \cdots \! \int f_x(y) \exp \left[\sum_{\kappa=1}^{k} (x_\kappa + iy_\kappa)t_\kappa \right] dv_y$$

which we can set up for all points $x \subset S$ is actually independent of x. We denote it by $\varphi(t)$.

By Plancherel's formulas we have, for $x \subset S$,

$$f_x(y) \sim \frac{1}{(2\pi)^k} \int_{-\infty}^{\infty} \!\! \cdots \! \int \varphi(t) \exp \left[- \sum_{\kappa=1}^{k} (x_\kappa + iy_\kappa)t_\kappa \right] dv_t \tag{6}$$

and

$$\int_{-\infty}^{\infty} \!\! \cdots \! \int | f_x(y) |^2 dv_y = \frac{1}{(2\pi)^k} \int_{-\infty}^{\infty} \!\! \cdots \! \int | \varphi(t) |^2 \exp \left[-2 \sum_{\kappa=1}^{k} x_\kappa t_\kappa \right] dv_t . \tag{7}$$

Thus, by (2),

$$\int_{-\infty}^{\infty}\cdots\int |\varphi(t)|^2 \exp\left(-2\sum_{\kappa=1}^{k} x_\kappa t_\kappa\right) dv_t \leqq (2\pi)^k K \tag{8}$$

for $x \subset S$. The function in (8) is the limit, as $a \to \infty$, of

$$\int_{-a}^{a}\cdots\int_{-a}^{a} |\varphi(t)|^2 \exp\left(-2\sum_{\kappa=1}^{k} x_\kappa t_\kappa\right) dv_t.$$

The latter function is convex along any straight line in R_k. Therefore it is $\leqq (2\pi)^k K$ for all $x \subset \tilde{S}$. We hence conclude that (8) holds for all $x \subset \tilde{S}$. Therefore we can construct the analytic function

$$g_x(y) \equiv g(z) = \frac{1}{(2\pi)^k}\int_{-\infty}^{\infty}\cdots\int \varphi(t)\exp\left[-\sum_{\kappa=1}^{k}(x_\kappa + iy_\kappa)t_\kappa\right] dv_t, \tag{9}$$

the integral converging absolutely and uniformly in every closed tube in the interior of $\tilde{T}$. Also

$$\int_{-\infty}^{\infty}\cdots\int |g_x(y)|^2 dv_y \leqq K \qquad \text{for } x \subset \tilde{S}.$$

The next step is to identify the functions $f(z)$ and $g(z)$. Comparing (6) and (9), we easily see that they are identical if some point x^0 of S is an inner point of $\tilde{S}$. Otherwise, if S lies on the boundary of $\tilde{S}$, there exist $k + 1$ points of S on the boundary of $\tilde{S}$ such that the rectilinear simplex with these points as vertices is k-dimensional and part of $\tilde{S}$.

Any non-singular linear transformation

$$Az_\kappa = a_{\kappa 1} z_1 + \cdots + a_{\kappa k} z_k$$

with real coefficients $a_{\kappa\lambda}$ carries a tube into a tube, and the basis of the one tube into the basis of the other. In particular, we can choose the transformation in such a way that (after the transformation) our simplex contains the basis of a tube

$$x_\kappa^0 < x_\kappa \leqq x_\kappa^0 + \eta, \qquad -\infty < y_\kappa < \infty, \tag{10}$$

the point x_κ^0 being a point of S. Our function $g(z)$ is analytic inside (10) and defined by (9), whereas $f(z)$ is analytic in a neighborhood of (1) and defined by (6) on (1), the function $\varphi(t)$ being the same in both cases. As in the case $k = 1$, we now readily conclude

$$\lim_{x_\kappa \to x_\kappa^0} \int_{-\infty}^{\infty}\cdots\int |f_{x^0}(t) - g_x(t)|^2 dv_t = 0.^4 \tag{11}$$

But the function $f_x(y)$ is continuous in a neighborhood of (1), hence by (11), for every finite $a > 0$,

$$\lim_{x_\kappa \to x_\kappa^0} \int_{-3a}^{3a}\cdots\int_{-3a}^{3a} |\delta_x(t)| dv_t = 0, \tag{12}$$

[4] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Oxford, 1937, p. 130, Theorem 97.

where

$$\delta(z) = f(z) - g(z)$$

is defined in the point set

$$x_\kappa^0 < x_\kappa \leqq x_\kappa^0 + \eta, \qquad 0 \leqq | y_\kappa | \leqq 3a,$$

for η sufficiently small (depending on a). For $0 < \epsilon < a$, the function

$$\delta^\epsilon(z) = \frac{1}{\epsilon^k} \int_0^\epsilon \cdots \int^\epsilon \delta(x_1 + iy_1 + iu_1, \cdots, x_k + iy_k + iu_k)\, dv_u$$

is analytic and bounded in

(13) $$x_\kappa^0 < x_\kappa \leqq x_\kappa^0 + \eta, \qquad 0 \leqq | y_\kappa | \leqq 2a,$$

and on account of (12) it tends to 0 as $x_\kappa \to x_\kappa^0$. For any fixed numbers y_κ, η_κ such that

$$0 \leqq | y_\kappa | \leqq a, \qquad 0 \leqq \eta_\kappa \leqq \max (1, \eta),$$

the function

$$\lambda_\epsilon(w) = \delta^\epsilon(x_1^0 + iy_1 + \eta_1 w, \cdots, x_k^0 + iy_k + \eta_k w)$$

of the complex variable $w = u + iv$ is analytic and bounded in the rectangle

$$0 < u \leqq 1, \qquad -a \leqq v \leqq a$$

and tends to 0 as $u \to 0$. Hence $\lambda_\epsilon(w) = 0$. In particular $\lambda_\epsilon(1) = 0$, or

$$\delta^\epsilon(x_1^0 + \eta_1 + iy_1, \cdots, x_k^0 + \eta_k + iy_k) = 0$$

for all values (13). Thus $\delta^\epsilon(z)$ vanishes on a $2k$-dimensional set and is therefore identically zero. But $\delta(z) = \lim_{\epsilon \to 0} \delta^\epsilon(z)$, and hence $\delta(z) = 0$, or $f(z) = g(z)$.

This completes the proof of our theorem under the additional assumption (3). If $f(z)$ is bounded,

$$f^\sigma(z) = f(z) \exp [\sigma(z_1^2 + \cdots + z_k^2)]$$

will satisfy (3) for $0 < \sigma < 1$, and since, for (5),

$$\int_{-\infty}^{\infty} \cdots \int | f_x^\sigma(y) |^2 dv_y \leqq A(\alpha) \int_{-\infty}^{\infty} \cdots \int | f_x(y) |^2 dv_y$$

and

$$\lim_{\sigma \to 0} \int_{-\infty}^{\infty} \cdots \int | f_x^\sigma(y) |^2 dv_y = \int_{-\infty}^{\infty} \cdots \int | f_x(y) |^2 dv_y,$$

our theorem holds for $f^\sigma(z)$, and for $f(z)$ itself.

PRINCETON UNIVERSITY.

APPROXIMATION TO THE SOLUTION OF A NORMAL SYSTEM OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS

By W. C. Risselman

1. **Introduction.** This note is concerned with certain problems of approximation on a given finite interval $a \leqq t \leqq b$ to the solution of the system of ordinary equations

$$(1) \qquad \begin{cases} \dfrac{dx_i}{dt} = \theta_{i1}(t)x_1 + \cdots + \theta_{im}(t)x_m + \theta_i(t), \\ x_i(t_0) = c_i, \end{cases} \qquad (i = 1, \cdots, m),$$

where $a \leqq t_0 \leqq b$. Let there be given an infinite sequence of functions $\varphi_i(t)$ which are defined and linearly independent (in finite subsets) on (a, b). Let

$$y_{in_i}(t) = c_{i1}\varphi_1(t) + \cdots + c_{in_i}\varphi_{n_i}(t) \qquad (i = 1, \cdots, m).$$

One may consider the problem of approximating to the solution of the system (1) by means of a set of m linear combinations $y_{in_i}(t)$ satisfying the initial conditions so as to minimize the sum

$$(2) \qquad \int_a^b |y'_{1n_1} - \theta_{11}y_{1n_1} - \cdots - \theta_{1m}y_{mn_m} - \theta_1|^{r_1}\,dt + \cdots \\ + \int_a^b |y'_{mn_m} - \theta_{m1}y_{1n_1} - \cdots - \theta_{mm}y_{mn_m} - \theta_m|^{r_m}\,dt,$$

where the r_i are given constants > 0. Another problem is that of approximating to the solution by means of a set of m linear combinations $y_{in_i}(t)$ so as to minimize

$$(3) \qquad \sum_{i=1}^{m} a_i\,|x_i(t_0) - y_{in_i}(t_0)|^{r_{m+i}} + \int_a^b |y'_{1n_1} - \theta_{11}y_{1n_1} - \cdots - \theta_{1m}y_{mn_m} - \theta_1|^{r_1}\,dt \\ + \cdots + \int_a^b |y'_{mn_m} - \theta_{m1}y_{1n_1} - \cdots - \theta_{mm}y_{mn_m} - \theta_m|^{r_m}\,dt,$$

where the r_i and the a_i are given constants > 0. Under further suitable hypotheses regarding the functions involved in these problems, questions of existence and uniqueness of approximating sets of functions will be discussed. The problem of uniform convergence as $n_i \to \infty$ of the $y_{in_i}(t)$ to the $x_i(t)$ will be discussed only in case the y_{in_i} are polynomials of degree at most n_i in t.

A number of papers[1] have been written recently dealing with similar problems

Received October 7, 1937.

[1] W. H. McEwen, Trans. Amer. Math. Soc., vol. 33(1931), pp. 979–997; Bulletin Amer. Math. Soc., vol. 38(1932), pp. 887–894. For other references to the literature see these papers by McEwen.

of approximation to the solution of a system consisting of a single m-th order linear differential equation and linearly independent two-point boundary conditions. The present problems, which are more general in certain respects and in which the initial conditions may be applied in the interior of the interval in which the approximation is made, have not yet been studied.

2. Preliminary theorem on existence and uniqueness.

DEFINITION. The m sets $\rho_{1j}(t), \cdots, \rho_{sj}(t)$ $(j = 1, \cdots, m)$ of s functions each will be said to be *properly independent* on (a, b) in case each of the functions is defined and measurable on (a, b) and at least one of the expressions $c_1\rho_{1j}(t) + \cdots + c_s\rho_{sj}(t)$ $(j = 1, \cdots, m)$ is different from zero on a subset of (a, b) of positive measure if the c_i are not all zero. If there exists a set of c's not all zero such that all of the expressions vanish almost everywhere on (a, b), the m sets $\rho_{1j}(t), \cdots, \rho_{sj}(t)$ will be said to be essentially dependent on (a, b).

LEMMA. *Let there be given m sets of s functions each $\rho_{1j}(t), \cdots, \rho_{sj}(t)$ which are defined on (a, b) and which satisfy the following conditions*:

(a) *The functions $\rho_{1j}, \cdots, \rho_{sj}$ belong to the Lebesgue class $L^{r_j}(a, b)$ $(j = 1, \cdots, m)$, where the r_j are given numbers > 0.*

(b) *The m sets are properly independent on (a, b).*

Let

$$T = \int_a^b |c_1\rho_{11} + c_2\rho_{21} + \cdots + c_s\rho_{s1}|^{r_1}\,dt + \cdots + \int_a^b |c_1\rho_{1m} + c_2\rho_{2m} + \cdots + c_s\rho_{sm}|^{r_m}\,dt.$$

Then all the c_i for which $|c_i| \geqq 1$ satisfy $|c_i| \leqq (T/T_0)^{1/r_s}$, where T_0 is a constant > 0 and r_s is one of the r_i which is at least as small as any other r.

Proof. Let E denote the set of points $(c_1, \cdots, c_s)$ in s dimensions such that $|c_i| = 1$ for some i and $|c_i| \leqq 1$ for all $i = 1, 2, \cdots, s$. Since T is a continuous function of the c's, since the set E is closed, and since the sets of functions are properly independent, it follows that T has a minimum $T_0 > 0$ on E. Now let the c's be arbitrary. Suppose that the absolute values of the c's are not all < 1 and that c_k is one of the c's whose absolute value is at least as great as that of any other c. Then the coefficients in $(c_1\rho_{1j} + \cdots + c_s\rho_{sj})/c_k$ belong to E. Since $|c_k| \geqq 1$, it follows that $|c_k| \leqq (T/T_0)^{1/r_s}$. Therefore all c's for which $|c_i| \geqq 1$ satisfy $|c_i| \leqq (T/T_0)^{1/r_s}$ and the proof of the lemma is complete.

Next the following preliminary theorem will be proved.

THEOREM A. *If the functions $\rho_{1j}(t), \cdots, \rho_{sj}(t)$ and $\theta_j(t)$ are defined on (a, b) and belong to the class $L^{r_j}(a, b)$ $(j = 1, \cdots, m)$, and if the m sets $\rho_{1j}, \cdots, \rho_{sj}$ are properly independent on (a, b), then there exists at least one set of values of the c's for which*

$$S = \int_a^b |\theta_1 - c_1\rho_{11} - \cdots - c_s\rho_{s1}|^{r_1}\,dt + \cdots + \int_a^b |\theta_m - c_1\rho_{1m} - \cdots - c_s\rho_{sm}|^{r_m}\,dt$$

is a minimum. If all r_j satisfy $r_j > 1$, the set of minimizing functions is unique.

Proof. If the m sets $\rho_{1j}, \cdots, \rho_{sj}, \theta_j$ of $(s + 1)$ functions each are essentially dependent on (a, b), there is a unique set of c's for which $S = 0$. Suppose that this is not the case. Then by virtue of the lemma all c_i for which $|c_i| \geqq 1$ satisfy $|c_i| \leqq (S/S_0)^{1/r_s}$, where $S_0 > 0$ is the minimum of

$$S' = \int_a^b |c_0\theta_1 - c_1\rho_{11} - \cdots - c_s\rho_{s1}|^{r_1}\,dt + \cdots$$
$$+ \int_a^b |c_0\theta_m - c_1\rho_{1m} - \cdots - c_s\rho_{sm}|^{r_m}\,dt$$

on the closed set in $(s + 1)$ dimensions for which $|c_i| = 1$ for some i and $|c_i| \leqq 1$ for all $i = 0, 1, \cdots, s$. Since S is a continuous function of the c's, and since the values of the c_i for which S is not greater than some specified number lie in a closed region, it follows that a set of minimizing functions exists. If all of the r_i are greater than unity, there is a unique set of minimizing functions, for the argument that if there were two sets of minimizing functions $\{x_{is}(t)\}$ and $\{y_{is}(t)\}$, then the set $\{(x_{is} + y_{is})/2\}$ would give a smaller value to S applies here.

3. **Existence and uniqueness of minimizing functions corresponding to the differential system.** In this section the following assumptions will be made regarding the functions $\theta_{ij}(t)$ and $\theta_i(t)$:

(a) Each of these functions is summable on (a, b).

(b) If any of the r_i $(i = 1, \cdots, m)$ are greater than unity, the functions $\theta_{ij}(t)$ and $\theta_i(t)$ belong to the class $L^{r_i}(a, b)$.

Besides the conditions imposed in the introduction the functions $\varphi_i(t)$ will be assumed to satisfy the following conditions in this section:

(a) Each of these functions is absolutely continuous on (a, b).

(b) If any of the r_i are greater than unity, and if r_g denotes one of the r's which is at least as large as any other r, each function satisfies the condition that the sum $\sum |\varphi_i(t_\nu - h_\nu) - \varphi_i(t_\nu)|^{r_g} h_\nu^{1-r_g}$ taken over any finite or enumerable system of non-overlapping intervals on (a, b) is bounded.

Under the conditions imposed on the functions $\theta_{ij}(t)$ and $\theta_i(t)$ there exists a unique set of absolutely continuous functions $x_1(t), \cdots, x_m(t)$ which satisfies the system (1) almost everywhere on (a, b).[2] One may then consider the problem of approximating this solution. The hypothesis (b) regarding the functions $\varphi_i(t)$ is necessary in order that these functions may have derivatives belonging to the class $L^{r_g}(a, b)$.[3]

In the discussion of the existence of a set of m combinations $y_{in_i}(t)$ satisfying

[2] W. M. Whyburn, Annals of Mathematics, (2), vol. 30(1928–29), pp. 31–38.

[3] See, for example, E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, pp. 384–386.

the initial conditions which minimize the sum (2) the following cases will be distinguished:

(a) The c_i are all zero and each function $\varphi_i(t)$ vanishes at $t = t_0$.

(b) The c_i are all zero but some of the φ's do not vanish at t_0.

(c) The c_i are not all zero.

Let

$$L_i(x_1, \cdots, x_m) = \frac{dx_i}{dt} - \theta_{i1}(t)x_1 - \cdots - \theta_{im}(t)x_m.$$

Then the sum (2) may be written as follows:

$$\int_a^b |\theta_1(t) - L_1(y_{1n_1}, \cdots, y_{mn_m})|^{r_1}\,dt + \cdots + \int_a^b |\theta_m(t) - L_m(y_{1n_1}, \cdots, y_{mn_m})|^{r_m}\,dt.$$

Let $l_{ijk}(t) = L_i(0, 0, \cdots, \varphi_k, \cdots, 0)$, where the notation means that the j-th argument is φ_k and all the other arguments are zero. Then

$$L_i(y_{1n_1}, \cdots, y_{mn_m}) = \sum_{j=1}^{m} \sum_{k=1}^{n_j} c_{jk} l_{ijk}(t).$$

Now, in case (a) the m sets of $n_1 + n_2 + \cdots + n_m$ functions each, which are obtained here, are properly independent. For if the $L_i(y_{1n_1}, \cdots, y_{mn_m})$ vanish simultaneously almost everywhere on (a, b), provided the c_{jk} are not all zero, then the homogeneous system $L_i(x_1, \cdots, x_m) = 0$, $x_i(t_0) = 0$ has a non-trivial solution on account of the linear independence of the φ_i. This is impossible. It follows from Theorem A that for each set of values of the n_i there exists a minimizing set of combinations and that, in case the r_i are greater than unity, this set is unique.

Case (b) will now be considered. Let $\varphi_k(t)$ be the first of the functions φ_i which does not vanish at t_0. If there are any values of i for which $n_i < k$, each of the φ's in the corresponding $y_{in_i}(t)$ vanishes at t_0. If there are values of i for which $n_i = k$, then $c_{in_i} = 0$ in the corresponding y_{in_i}. Consider the values of i for which $n_i > k$. Let

$$\psi_i(t) = \varphi_i(t) - \frac{\varphi_i(t_0)}{\varphi_k(t_0)}\varphi_k(t) \qquad (i = 1, \cdots, k-1, k+1, \cdots, n_i).$$

Then each of the functions $\psi_i(t)$ vanishes at $t = t_0$. From the linear independence of the φ_i it follows that the ψ_i are linearly independent on (a, b). For the values of i for which $n_i > k$ the approximation is to be made in terms of linear combinations of the ψ_i. Thus it is seen that in this case either the y_{in_i} are identically zero or else, as in case (a), the approximation is to be made in terms of m properly independent sets of functions, and therefore by Theorem A there exists a minimizing set of combinations which is unique in case the r_i exceed unity.

In the consideration of case (c) it is to be noticed that if for each i for which $c_i \neq 0$ there exists a $\varphi_k(t)$ with $k \leqq n_i$ which does not vanish at t_0, $c_i\varphi_k(t)/\varphi_k(t_0)$

is an admissible combination which satisfies the initial condition. Then the problem is that of approximating the set $\theta_i(t) - L_i(c_1\varphi_k(t)/\varphi_k(t_0), \cdots, c_m\varphi_k(t)/\varphi_k(t_0))$ with a set $L_i(Y_{1n_1}, \cdots, Y_{mn_m})$, where the Y_{in_i} vanish at t_0. If there is a value of i such that $c_i \neq 0$ and all the φ's with subscripts $\leqq n_i$ vanish at t_0, the initial conditions cannot be satisfied.

Next it will be shown that for each set of values of the n_i there exists a set y_{in_i} which minimizes the sum (3). This set is unique in case the r_i $(i = 1, \cdots, 2m)$ are greater than unity. If for a given set n_i each of the φ_i in question vanishes at t_0, the set that minimizes the sum (3) is the same set that minimizes the sum (2) and satisfies $y_{in_i}(t_0) = 0$. If some of the φ_i do not vanish at t_0, let φ_k be the first function in the sequence which does not vanish there. Then for all values of i for which $n_i \geqq k$ one has

$$y_{in_i}(t) = c_{i1}\psi_1(t) + \cdots + c_{ik-1}\psi_{k-1}(t) + c_{ik+1}\psi_{k+1}(t) + \cdots + c_{in_i}\psi_{n_i}(t) + \frac{b_i}{\varphi_k(t_0)}\varphi_k(t),$$

where the ψ_i have the same meaning as before. One has $y_{in_i}(t_0) = b_i$. On account of the initial conditions imposed, one need consider only values of the b_i for which $|b_i| \leqq$ some bound M. For each value of i one may write

$$L_i(y_{1n_1}, \cdots, y_{mn_m}) = L_i\left(\frac{b_1}{\varphi_k(t_0)}\varphi_k(t), \cdots, \frac{b_m}{\varphi_k(t_0)}\varphi_k(t)\right) + L_i(Y_{1n_1}, \cdots, Y_{mn_m}),$$

where each of the functions in the Y_{in_i} vanishes at t_0 and the b_i are zero in case $k > n_i$. It follows from the previous work that for each fixed set of values of the b_i a minimum $m(b_1, \cdots, b_m)$ of (3) exists. But this function is a continuous function of the b_i and one needs to consider only a closed region. In this region a minimum of $m(b_1, \cdots, b_m)$ exists and this is the minimum of (3).

4. **Convergence in case $r_1, \cdots, r_m$ are $\geqq 1$.** In the study of the problem of convergence it will be assumed at first that the functions $\theta_{ij}(t)$ and $\theta_i(t)$ are continuous on (a, b) and that the approximations to the $x_i(t)$ are to be made by polynomials $P_{in_i}(t)$ of degree at most n_i. Use will be made of the following auxiliary theorem:

THEOREM B. *Let K be a given real number $\geqq 0$; let $a_1, \cdots, a_m$ and $r_{m+1}, \cdots, r_{2m}$ be given real numbers > 0, and let $r_1, \cdots, r_m$ be given real numbers $\geqq 1$. Then if $u_1(t), \cdots, u_m(t)$ are any set of m functions which are defined and have continuous first derivatives on $a \leqq t \leqq b$ and which satisfy the relation*

$$S = \int_a^b |u_1' - \theta_{11}u_1 - \cdots - \theta_{1m}u_m - \theta_1|^{r_1}\,dt + \cdots + \int_a^b |u_m' - \theta_{m1}u_1 - \cdots - \theta_{mm}u_m - \theta_m|^{r_m}\,dt + \sum_{i=1}^m a_i\,|x_i(t_0) - u_i(t_0)|^{r_{m+i}} \leqq K,$$

there exist constants M_{ij} $(i = 1, \cdots, m;\ j = 1, \cdots, 2m)$ *independent of the functions* $u_i(t)$ *such that*

$$|x_i(t) - u_i(t)| \leqq M_{i1}K^{1/r_1} + \cdots + M_{i,2m}K^{1/r_{2m}} \text{ on } (a, b) \qquad (i = 1, \cdots, m).$$

Let the functions $\eta_i(t)$ be defined on (a, b) by the equations

$$\frac{du_i}{dt} = \theta_{i1}(t)u_1 + \cdots + \theta_{im}(t)u_m + \theta_i(t) + \eta_i(t) \qquad (i = 1, \cdots, m). \tag{4}$$

Let $q_i(t) = u_i(t) - x_i(t)$. By subtraction one obtains from equations (1) and (4)

$$\frac{dq_i}{dt} = \theta_{i1}(t)q_1 + \cdots + \theta_{im}(t)u_m + \eta_i(t) \qquad (i = 1, \cdots, m). \tag{5}$$

By virtue of the relation which the u_i satisfy by hypothesis

$$|q_i(t_0)| = |x_i(t_0) - u_i(t_0)| \leqq \left(\frac{K}{a_i}\right)^{1/r_{m+i}}. \tag{6}$$

Let the functions $\varphi_{ij}(t)$ be the fundamental set for the system

$$\frac{dx_i}{dt} = \theta_{i1}(t)x_1 + \cdots + \theta_{im}(t)x_m \qquad (i = 1, \cdots, m)$$

which satisfy the initial conditions $\varphi_{ii}(t_0) = 1$, $\varphi_{ij}(t_0) = 0$ if $i \neq j$, where the notation means that one has a solution if a fixed value is assigned to j. Let $D(t)$ denote the determinant $|\varphi_{ij}(t)|$ and let $w_{ij}(t)$ denote the cofactor of $\varphi_{ij}(t)$ in this determinant. Then

$$D(t) = \exp\left[\int_{t_0}^{t} \sum_{i=1}^{m} \theta_{ii}(t)\,dt\right]$$

is continuous and does not vanish on (a, b).

The solution of the system (1) is given by

$$\begin{aligned} x_i(t) &= c_1\varphi_{i1}(t) + \cdots + c_m\varphi_{im}(t) \\ &\quad + \varphi_{i1}(t)\int_{t_0}^{t} \frac{\sum_{j=1}^{m} \theta_j(t)w_{j1}(t)}{D(t)}\,dt + \cdots + \varphi_{im}(t)\int_{t_0}^{t} \frac{\sum_{j=1}^{m} \theta_j(t)w_{jm}(t)}{D(t)}\,dt. \end{aligned} \tag{7}$$

It follows that the $q_i(t)$ satisfy the equations

$$\begin{aligned} q_i(t) &= q_i(t_0)\varphi_{i1}(t) + \cdots + q_m(t_0)\varphi_{im}(t) \\ &\quad + \varphi_{i1}(t)\int_{t_0}^{t}\left[\frac{\eta_1 w_{11}}{D} + \cdots + \frac{\eta_m w_{m1}}{D}\right]dt + \cdots \\ &\quad + \varphi_{im}(t)\int_{t_0}^{t}\left[\frac{\eta_1 w_{1m}}{D} + \cdots + \frac{\eta_m w_{mm}}{D}\right]dt. \end{aligned}$$

By a straightforward process in which use is made of the Hölder inequality

$$\int_a^b |\eta_j(t)|\,dt \leqq (b-a)(r_j-1)/r_j\left[\int_a^b |\eta_j(t)|^{r_j}\,dt\right]^{1/r_j},$$

it is found that

$$(8)\quad \begin{aligned} |q_i(t)| \leqq M_{i1}\left[\int_a^b |\eta_1|^{r_1}\,dt\right]^{1/r_1} + \cdots + M_{im}\left[\int_a^b |\eta_m|^{r_m}\,dt\right]^{1/r_m} \\ + |q_i(t_0)\varphi_{i1}(t)| + \cdots + |q_m(t_0)\varphi_{im}(t)|, \end{aligned}$$

where M_{ij} is the maximum on (a, b) of $(b-a)^{(r_j-1)/r_j}\sum_{k=1}^{m} A_{jk}\,|\varphi_{ik}(t)|$, and A_{jk} is the maximum on (a, b) of the continuous function $|w_{jk}D^{-1}|$. But it follows directly from the hypothesis that

$$\left[\int_a^b |\eta_j|^{r_j}\,dt\right]^{1/r_j} \leqq K^{1/r_j} \qquad (j = 1, \cdots, m).$$

Let N_{ik} equal the maximum of $|\varphi_{ik}(t)|$ on (a, b). Then from (6)

$$|q_k(t_0)\varphi_{ik}(t)| \leqq (Ka_k^{-1})^{1/r_{m+k}}N_{ik} \qquad (i = 1, \cdots, m; k = 1, \cdots, m).$$

In the last relation let the coefficient of $K^{1/r_{m+k}}$ be denoted by M_{ij} $(i = 1, \cdots, m; j = m+1, \cdots, 2m)$. Taking (8) into consideration, one may now write

$$|q_i(t)| = |x_i(t) - u_i(t)| \leqq M_{i1}K^{1/r_1} + \cdots + M_{i2m}K^{1/r_{2m}}$$

on (a, b).

Hereafter a set of polynomials $\{P_{in_i}\}$ which satisfies the initial conditions and minimizes the expression (2) will be called *an approximating set of the first kind*, and a set $\{P_{in_i}\}$ which minimizes (3) will be called *an approximating set of the second kind*.

A way has now been prepared for the discussion of convergence problems if $r_1, \cdots, r_m$ are $\geqq 1$. Since the $x_i(t)$ are continuous and have continuous first derivatives on (a, b), it follows as a corollary to Theorem B of McEwen's paper[4] that there exists a set of polynomials $\{P_{in_i}\}$ satisfying the initial conditions, corresponding to each set of positive integers $\{n_i\}$ such that

$$|x_i(t) - P_{in_i}(t)| \leqq \epsilon_{in_i}; \qquad |x_i'(t) - P_{in_i}'(t)| \leqq \epsilon_{in_i},$$

where $\lim_{n_i\to\infty} \epsilon_{in_i} = 0$.

Let $f_i(t) = x_i(t) - P_{in_i}(t)$. Then the $f_i(t)$ satisfy the system

$$\begin{cases} \dfrac{df_i}{dt} = \theta_{i1}(t)f_1 + \cdots + \theta_{im}(t)f_m + y_i(t), \\ f_i(t_0) = 0, \end{cases} \qquad (i = 1, \cdots, m),$$

[4] McEwen, Transactions, loc. cit., p. 983.

where the $y_i(t)$ satisfy $|y_i(t)| \leqq (m+1)D_i\epsilon_g$. Here D_i is an upper bound of the bounded functions $1, |\theta_{i1}|, \cdots, |\theta_{im}|$, and ϵ_g is the largest of the numbers ϵ_{in_i}. Now the $P_{in_i}(t)$ have the property that they make the sum S of Theorem B equal to

$$\int_a^b |y_1(t)|^{r_1}\,dt + \cdots + \int_a^b |y_m(t)|^{r_m}\,dt.$$

This sum is less than or equal to

$$\delta_{(n_i)} = m(b-a)[(m+1)D]^{r_g}\epsilon_g^{r_s} \qquad \text{if} \qquad \epsilon_g < 1,$$

where D is the largest of the numbers D_i, r_s is the smallest of the r_i, and r_g is the largest of the r_i. Since an approximating set of the first kind $\{P_{in_i}\}$ makes the sum S less than or equal to $\delta_{(n_i)}$ if the n_i are sufficiently large, it follows from Theorem B that for such values of the n_i

$$|x_i(t) - P_{in_i}(t)| \leqq M_{i1}\delta_{(n_i)}^{1/r_1} + \cdots + M_{i2m}\delta_{(n_i)}^{1/r_{2m}} \qquad (i = 1, \cdots, m).$$

Since $\delta_{(n_i)} \to 0$ as $n_i \to \infty$, one has the following result:

THEOREM I. *If $\{P_{in_i}(t)\}$ denotes an approximating set of the first kind, if $\epsilon > 0$ is arbitrarily assigned, and if the functions $\theta_{ij}(t)$ and $\theta_i(t)$ are continuous on (a, b), then there exists an integer N_ϵ such that, if the n_i are $\geqq N_\epsilon$, then $|x_i(t) - P_{in_i}(t)| < \epsilon$ for all values of t on (a, b).*

Re-examination of the proof of Theorem I will show that one immediately has the following:

THEOREM II. *If $\{P_{in_i}(t)\}$ denotes an approximating set of the second kind, if $\epsilon > 0$ is arbitrarily assigned, and if the functions $\theta_{ij}(t)$ and $\theta_i(t)$ are continuous on (a, b), there exists an integer N_ϵ such that, if the n_i are $\geqq N_\epsilon$, then $|x_i(t) - P_{in_i}(t)| < \epsilon$ for all values of t on (a, b).*

The hypotheses on the functions $\theta_{ij}(t)$ and $\theta_i(t)$ will now be made more restrictive. One may prove the following theorem:

THEOREM III. *If each of the functions $\theta_{ij}(t)$ and $\theta_i(t)$ satisfies an ordinary Lipschitz condition on (a, b), and if $\{P_{in_i}(t)\}$ denotes an approximating set of the first kind, then there exists a constant G independent of the n_i such that for every set of positive integers $\{n_i\}$, $|x_i(t) - P_{in_i}(t)| \leqq G/n_s^r$ on (a, b), where r is the smallest of the numbers r_i/r_j $(i = 1, \cdots, m; j = 1, \cdots, m)$, and n_s is the smallest integer n_i.*

Since each of the functions $\theta_i(t)$ and $\theta_{ij}(t)$ satisfies an ordinary Lipschitz condition on (a, b), the functions $x_i'(t)$ satisfy a Lipschitz condition on this interval. Then it is a corollary to Theorem D of McEwen's paper[5] that for each set of positive integers n_i there exists a set of polynomials $\{P_{in_i}(t)\}$ satisfying the initial conditions such that

$$|x_i(t) - P_{in_i}(t)| \leqq \frac{C}{n_i}, \qquad |x_i'(t) - P_{in_i}'(t)| \leqq \frac{C}{n_i}, \qquad (i = 1, \cdots, m)$$

[5] McEwen, Transactions, loc. cit., p. 985.

for all values of t on (a, b), where C is a constant independent of the n_i. The rest of the proof is similar to the proof of Theorem I. The expressions C/n_i now take the place of the ϵ_{in_i} used in the proof of that theorem. Among the points to be noted are the facts that C/n_i is less than unity if n_i is sufficiently large, that $\delta_{(n_i)}$ is less than unity if n_s is sufficiently large, and that one may choose each of the numbers $r_{m+1}, \cdots, r_{2m}$ less than the r_i in the expression to be minimized.

In like manner one obtains the following result:

THEOREM IV. *If each of the functions $\theta_{ij}(t)$ and $\theta_i(t)$ satisfies an ordinary Lipschitz condition on (a, b), and if $\{P_{in_i}(t)\}$ denotes an approximating set of the second kind, then there exists a constant H independent of the n_i such that for every set of positive integers $\{n_i\}$, $| x_i(t) - P_{in_i}(t) | \leqq H/n_s^r$ on (a, b), where r is the smallest of the numbers r_i/r_j $(i = 1, \cdots, m; j = 1, \cdots, 2m)$ and n_s is the smallest integer n_i.*

A further result obtained by the same method of proof is the following:

THEOREM V. *If each of the functions $\theta_{ij}(t)$ and $\theta_i(t)$ has a k-th derivative satisfying an ordinary Lipschitz condition, and if $\{P_{in_i}(t)\}$ denotes either an approximating set of the first kind or an approximating set of the second kind, then there exists a constant $\mathfrak{L}$ independent of the n_i such that for every set of positive integers $\{n_i\}$, $| x_i(t) - P_{in_i}(t) | \leqq \mathfrak{L}/n_s^{(k+1)r}$ on (a, b), where r and n_s have the same meaning as they have in the corresponding cases in Theorems* III *and* IV.

In the proof of this theorem one notices that it follows from the hypotheses that $x_i(t)$ has a $(k + 1)$-th derivative satisfying an ordinary Lipschitz condition. Use is made of the remark made after Theorem D of McEwen's paper.[6]

5. **Convergence if the r_i are < 1.** In the treatment in the preceding section of the convergence problem use was made of the Hölder inequality, a relation which is not applicable if some or all of the r_i $(i = 1, \cdots, m)$ are less than unity. Consequently a different method must be used in this section. This method, which is an extension of the method which McEwen used[7] to treat the corresponding case in his problems, is not restricted in application to cases in which the r_i are < 1, but the bound which it assigns to the errors in case the r_i are $\geqq 1$ is not as good as that obtained by the preceding treatment. It does, however, prove the convergence of the set $\{P'_{in_i}\}$ to the set $\{x'_i\}$ as well as the convergence of the set $\{P_{in_i}\}$ to the set $\{x_i\}$ as the $n_i \to \infty$.

In the proof of the convergence theorem use is made of the following result:

THEOREM C. *If any set of polynomials $\{\pi_{in_i}(t)\}$ satisfies the initial conditions $\pi_{in_i}(t_0) = d_i$, if the θ_{ij} and the θ_i are continuous on (a, b), and if δ_g is one of the numbers δ_i which is at least as large as any other δ_i, where $\delta_i = \max | L_i(\pi_{1n_1}, \cdots, \pi_{mn_m}) |$ on $a \leqq t \leqq b$, then*

$$\text{(a)} \qquad | \pi_{in_i}^{(k)} | \leqq C\left(\delta_g + \sum_{i=1}^{m} | d_i |\right) \qquad (k = 0, 1);$$

[6] McEwen, Transactions, loc. cit., p. 986.

[7] McEwen, Bulletin, loc. cit., pp. 891–894.

and if the functions $\theta_{ij}(t)$ have bounded first derivatives

$$\text{(b)} \qquad \left| \frac{d}{dt} L_i(\pi_{1n_1}, \cdots, \pi_{mn_m}) \right| \leqq Dn_i^2 \left(\delta_g + \sum_{i=1}^{m} |d_i| \right) \qquad (i = 1, \cdots, m)$$

for all values of t on (a, b), where C and D are constants independent of the n_i and of the coefficients in the π_{in_i}.

The proof of this theorem is similar to the proof of McEwen's Auxiliary Theorem.[8] Let $Z_i(t) = L_i(\pi_{1n_1}, \cdots, \pi_{mn_m})$. Then the set $\{\pi_{in_i}\}$ is the solution of the system

$$L_i(x_1, \cdots, x_m) = Z_i(t), \qquad x_i(t_0) = d_i, \qquad (i = 1, \cdots, m).$$

Therefore π_{in_i} is given by the formula (7). The remainder of the proof of statement (a) can be supplied readily. In the proof of conclusion (b) use is made of Markoff's theorem on the derivative of a polynomial.

The following theorem on convergence can now be proved.

THEOREM VI. *If the functions $\theta_i(t)$ are continuous on (a, b), if each of the functions $\theta_{ij}(t)$ has a bounded first derivative on (a, b), and if $\{P_{in_i}(t)\}$ denotes either an approximating set of the first kind or an approximating set of the second kind, then there exists a constant C' independent of the n_i such that for every set of positive integers $\{n_i\}$*

$$|x_i^{(k)}(t) - P_{in_i}^{(k)}(t)| \leqq C' n_g^{2/r_s} \epsilon_g^s \qquad (k = 0, 1) \text{ on } (a, b),$$

where n_g is the largest of the n_i, r_s is the smallest of the r_i $(i = 1, \cdots, m)$, s is the smallest of the numbers r_i/r_j $(i = 1, \cdots, m; j = i, \cdots, 2m)$, and ϵ_g is the largest of the numbers ϵ_{in_i}, where ϵ_{in_i} is an upper bound of the error with which $x_i^{(k)}(t)$ can be approximated by polynomials $p_{in_i}^{(k)}(t)$ $(k = 0, 1)$ of degree at most n_i satisfying the initial conditions.

The proof, which will be omitted, is similar to the corresponding proof of convergence given in McEwen's paper.[9]

ARIZONA STATE TEACHERS COLLEGE.

[8] McEwen, Bulletin, loc. cit., pp. 891 and 892.
[9] McEwen, Bulletin, loc. cit., pp. 892–894.

SOLUTIONS OF A DIFFERENTIAL EQUATION OF THE FIRST ORDER AND FIRST DEGREE IN THE VICINITY OF BRANCH POINTS OF THE SOLUTION

By Jesse Pierce

Introduction. In this paper a method previously developed[1] for solving systems of differential equations about an ordinary point is applied to the equation

$$2x\frac{dx}{dt} = \sum_{h=0}^{\infty} f_h(t)x^h \qquad (f_0(t) \neq 0), \tag{1}$$

which has a singularity at $x = 0$.

The coefficients $f_h(t)$, which for convenience will be represented by f_h, are assumed to satisfy the following conditions:

I. The functions f_h $(h = 1, 2, \cdots)$ are real and are dominated on the interval (t_0, t_1) by a function f of t. The functions f_h $(h = 0, 1, \cdots)$ are integrable (Riemann) and their only points of discontinuity belong to a set E of measure zero.

II. The function f is integrable (Riemann) on the interval (t_0, t_1) and is equal to unity for all values of $t < t_0$. The points of discontinuity of f belong to the set E. The definite integral of f on the interval (t_0', t_0) where $t_0' \leqq t_0$ is represented by

$$\int_{t_0'}^{t_0} f\,dt = t_0 - t_0' = c \geqq 0. \tag{2}$$

III. The function f_0 is real and satisfies the inequalities

$$\begin{cases} 0 \leqq \displaystyle\int_{t_0}^{t} f_0\,dt \leqq \int_{t_0}^{t} f\,dt, \\[2ex] \dfrac{1}{\displaystyle\int_{t_0}^{t} f_0\,dt + c} \leqq \dfrac{A^2}{\displaystyle\int_{t_0}^{t} f\,dt + c} \equiv \dfrac{A^2}{F^2}, \end{cases} \tag{3}$$

where A is a real constant satisfying the inequality $A \geqq 1$ and the zeros of the function F belong to the set E.

1. **Formal solutions of the differential equation** (1). The transformation

$$x = \sum_{h=1}^{\infty} y_h K^h, \tag{4}$$

Received December 15, 1937; in revised form, October 7, 1938.

[1] J. Pierce, *Solutions of systems of differential equations in terms of infinite series of definite integrals*, this Journal, vol. 3(1937), pp. 616–622.

where K is an arbitrary parameter, reduces equation (1) to

$$\sum_{h=1}^{\infty} 2K^{h+1} \sum_{p=1}^{h} y_p y'_{h-p+1} = f_0 + \sum_{h=2}^{\infty} K^{h-1} \sum_{q=1}^{h-1} f_q y_{i_1} y_{i_2} \cdots y_{i_q} \tag{5}$$

$$(i_1 + i_2 + \cdots + i_q = h - 1),$$

where $y'_i = \dfrac{dy_i}{dt}$.

A formal solution of the differential equation (5) can be found by solving the following system of linear differential equations and then replacing K by unity:

$$\begin{cases} 2y_1 y'_1 = f_0, \\ 2\sum_{p=1}^{h} y_p y'_{h-p+1} = \sum_{q=1}^{h-1} f_q y_{i_1} y_{i_2} \cdots y_{i_q} \\ \qquad (i_1 + i_2 + \cdots + i_q = h - 1;\ h = 2, 3, \cdots). \end{cases} \tag{6}$$

Equations (6) are obtained by equating the coefficient of K^{h+1} on the left of (5) to the coefficient of K^{h-1} on the right.

The first differential equation in (6) has the solution

$$y_1^2 = \int_{t_0}^{t} f_0\, dt + c, \tag{7}$$

where c is defined by (2). Let η be a definite square root of the right member of (7). Since there are two choices for η, there are two distinct values for y_1, one of which will be represented by

$$y_1 = \eta. \tag{8}$$

The other differential equations in (7) can be solved, sequentially, in the form

$$\begin{cases} y_2 = \dfrac{1}{2\eta} \displaystyle\int_{t_0}^{t} f_1 \eta\, dt, \\ y_h = \dfrac{1}{2\eta} \displaystyle\int_{t_0}^{t} \left[\sum_{q=1}^{h-1} f_q y_{i_1} y_{i_2} \cdots y_{i_q}\right] dt - \dfrac{1}{2\eta} \sum_{p=2}^{h-1} y_p y_{h-p+1} \qquad (h = 3, 4, \cdots). \end{cases} \tag{9}$$

Hence two formal solutions of the differential equation (1) are represented by the series

$$x = \sum_{h=1}^{\infty} y_h \qquad (y_1 = \eta). \tag{10}$$

In order to obtain a satisfactory dominating series, consider the differential equation

$$\begin{aligned} \frac{d}{dt}[A^2FX] = \frac{A^4 f}{1 - X} + \left[\frac{A^4F^2X}{(1 - X)^2} + 2A^2FX - \frac{A^4F^2}{2}\right]\frac{dX}{dt} \\ + [X - A^2F]\left[\frac{dX}{dt} - \frac{A^2 f}{2F}\right]. \end{aligned} \tag{11}$$

When the right member of (11) is expanded in a power series in X, it can be written in the form

$$\frac{d}{dt}[A^2FX] = \sum_{h=0}^{\infty} fA^4X^h + \left[\sum_{h=1}^{\infty} A^4F^2hX^h + 2A^2FX - \frac{A^4F^2}{2}\right]\frac{dX}{dt} + [X - A^2F]\left[\frac{dX}{dt} - \frac{A^2f}{2F}\right]. \tag{12}$$

The transformation

$$X = A^2FK + \sum_{h=2}^{\infty} Y_hK^h = \sum_{h=1}^{\infty} Y_hK^h \qquad (Y_1 = A^2F) \tag{13}$$

reduces the differential equation (12) to

$$\begin{aligned}\sum_{h=1}^{\infty} K^h \frac{d}{dt}(A^2FY_h) &= A^4f + \sum_{h=2}^{\infty} K^{h-1} \sum_{q=1}^{h-1} A^4fY_{i_1}Y_{i_2}\cdots Y_{i_q} + \sum_{h=3}^{\infty} K^{h+1} \sum_{p=2}^{h-1} Y_pY'_{h-p+1}\\ &+ \sum_{h=2}^{\infty} K^h\left[2A^2F\sum_{p=1}^{h-1} Y_pY'_{h-p} - \frac{A^4F^2}{2K}Y'_{h-1}\right]\\ &+ \sum_{h=2}^{\infty} K^hA^4F^2\sum_{p=1}^{h-1} Y_pY'_{h-p} + \sum_{h=3}^{\infty} K^hA^4F^2G_h(Y_1, \cdots, Y_{h-2}, Y'_1, \cdots, Y'_{h-2})\\ &+ \frac{(K-1)A^2f}{2F}\sum_{h=2}^{\infty} K^hY_h + (K-1)A^2F\sum_{h=2}^{\infty} K^hY'_h + \frac{(K-1)^2A^4f}{2},\end{aligned} \tag{14}$$

where the function $G_h(Y_1, \cdots, Y_{h-2}, Y'_1, \cdots, Y'_{h-2})$ is the coefficient of K^h in the expansion of

$$\left\{\sum_{h=2}^{\infty} A^4F^2h\left[\sum_{k=1}^{\infty} K^kY_k\right]^h\right\}\left\{\sum_{h=1}^{\infty} K^hY'_h\right\} \tag{15}$$

as a power series in K. It is clear that the function G_h is a polynomial in the Y_i and Y'_j $(i, j = 1, \cdots, h-2)$ with positive coefficients.

A formal solution of the differential equation (14) can be found by replacing K by unity and solving, sequentially, the system

$$\left\{\begin{aligned}\frac{d}{dt}(A^2FY_1) &= A^4f,\\ \frac{d}{dt}(A^2FY_2) &= A^4fY_1 + \left[2A^2FY_1 - \frac{A^4F^2}{2}\right]Y'_1,\\ \frac{d}{dt}(A^2FY_h) &= \sum_{q=1}^{h-1} A^4fY_{i_1}Y_{i_2}\cdots Y_{i_q} + \sum_{p=2}^{h-1} Y_pY'_{h-p+1}\\ &+ 2A^2F\sum_{p=1}^{h-1} Y_pY'_{h-p} - \frac{A^4F^2}{2}Y'_{h-1}\\ &+ A^4F^2\sum_{p=1}^{h-2} Y_pY'_{h-p-1} + A^4F^2G_{h-1}(Y_1, \cdots, Y_{h-3}, Y'_1, \cdots, Y'_{h-3})\\ &\qquad (i_1 + \cdots + i_q = h-1;\ h = 3, 4, \cdots).\end{aligned}\right. \tag{16}$$

Equations (16) are obtained by equating the coefficient of K^h in the left member of (14) to the coefficient of K^h in the first sum in the right member plus the coefficient of K^{h+1} in the second sum plus the coefficient of K^h in the third sum plus the coefficients of K^{h-1} in the fourth and fifth sums. A formal solution of the system of differential equations (16) is

$$
(17)\quad \begin{cases}
Y_1 = \dfrac{A^2}{F}\displaystyle\int_{t_0'}^{t} f\,dt = A^2F \equiv A^2H_1, \\[2ex]
Y_2 = \dfrac{A^4}{F}\displaystyle\int_{t_0'}^{t} [H_1 f + \{2FH_1 - \tfrac{1}{2}F^2\}\,H_1']\,dt = \tfrac{7}{6}\,A^4F^2 \equiv A^4H_2, \\[2ex]
Y_h = \dfrac{A^{2h}}{F}\displaystyle\int_{t_0'}^{t} \Big[fH_{i_1}H_{i_2}\cdots H_{i_q} + 2F\sum_{p=1}^{h-1} H_pH'_{h-p} - \tfrac{1}{2}F^2H'_{h-1} \\
\qquad + F^2\displaystyle\sum_{p=1}^{h-2} H_pH'_{h-p-1} + F^2G_{h-1}(H_1, \cdots, H_{h-3}, H_1', \cdots, H'_{h-3}) \Big]\,dt \\
\qquad + \dfrac{A^{2h}}{2F}\displaystyle\sum_{p=2}^{h-1} H_pH_{h-p+1} \\
\qquad = a_hA^{2h}F^h \equiv A^{2h}H_h \qquad\qquad (h = 3, 4, \cdots),
\end{cases}
$$

where the a_h are known constants. Hence the differential equation (11) has the formal solution

$$
(18)\qquad X = \sum_{h=1}^{\infty} A^{2h}H_h,
$$

which is a power series in A^2F.

The differential equation (11) has the general integral

$$
(19)\qquad \tfrac{3}{2}A^4F^2 - 1 - 2A^2FX - \tfrac{1}{2}A^4F^2X + 2A^2FX^2 + X^2 = C(1 - X)e^X,
$$

where C is an arbitrary parameter. If we replace C by -1 and expand e^X in a power series in X, equation (19) takes the form

$$
(20)\qquad \begin{aligned}
&[-2A^2FX + \tfrac{3}{2}A^4F^2 + \tfrac{1}{2}X^2] \\
&\qquad + \left[-\tfrac{1}{2}A^4F^2X + 2A^2FX^2 - \sum_{h=3}^{\infty}\frac{X^h}{(h-1)!} + \sum_{h=3}^{\infty}\frac{X^h}{h!} \right] = 0.
\end{aligned}
$$

It follows from the theory of implicit functions that equation (20) can be solved for X as a power series in A^2F in two and only two distinct ways.[2] In one of these series the coefficient of A^2F is unity and in the other the corresponding coefficient is three. As the coefficient of A^2F in (18) is unity, it follows that the series (18) is the same as the expanded form of (20), which has the coefficient of A^2F equal to unity. Hence there is a radius of convergence $r > 0$ for which the series (18) will converge when

$$
(21)\qquad A^2F \leqq r.
$$

[2] F. R. Moulton, *Differential Equations*, p. 85.

As

$$(22) \qquad F^2 = \int_{t_0}^{t} f\,dt + c = \int_{t_0'}^{t_0} f\,dt + \int_{t_0}^{t} f\,dt = \int_{t_0'}^{t} f\,dt,$$

it follows that to make $A^2F < r$ the interval (t_0', t) will, in general, have to be limited to certain values; say $t - t_0' < R$. The point t_0 lies between t_0' and t, hence the sum of the lengths of the intervals (t_0', t_0) and (t_0, t) must be less than R if the series (18) is to converge.

It follows from the inequalities (3) that

$$(23) \qquad \eta \leqq H_1 \leqq AH_1 .$$

The second equation in (17) can be written in the form

$$(24) \qquad Y_2 = A^2 \frac{A}{F} \int_{t_0'}^{t} [AH_1F + A\{2FH_1 - \tfrac{1}{2}F^2\}H_1']\,dt.$$

The quantity

$$(25) \qquad 2FH_1 - \tfrac{1}{2}F^2 = \tfrac{3}{2}F^2$$

is positive and as $AF^{-1} \geqq \eta^{-1}$ (by (3)) it follows from (23), (24), (17) and (9) that

$$(26) \qquad | y_2 | \leqq A^2H_2 .$$

By induction it is readily proved that

$$(27) \qquad | y_h | \leqq A^hH_h ,$$

and hence

$$(28) \qquad | y_h | \leqq A^{2h}H_h = Y_h \qquad (A \geqq 1).$$

Thus the series (10) will converge when the inequality $t - t_0' < R$ is satisfied.

When $t_0' = t_0$, then $c = 0$ and the initial value of each of the two distinct functions represented by the series (10) is zero.

2. **Proof of the existence of the derivative of the series** (10). Enclose the set E in a sequence of intervals $\{\delta_i\}$ the sum of whose lengths is less than ϵ, which is positive but arbitrarily small. Delete this sequence of intervals $\{\delta_i\}$ from the interval (t_0, t_1). It follows, sequentially, from (9) that every y_h is a continuous function on the deleted interval (t_0, t_1). It follows from (3) and (28) that the integrand in the definite integral defining y_h is bounded on the deleted interval (t_0, t_1). Hence every y_h has a derivative at every point on the deleted interval (t_0, t_1).

If we compare (16) with (6), it is clear that

$$(29) \qquad | y_h' | \leqq A^{2h}H_h' .$$

The series $\sum_{h=1}^{\infty} A^{2h} H_h'$ converges uniformly on the deleted interval (t_0, t_2), for which the series (18) converges, and hence the series $\sum_{h=1}^{\infty} y_h'$ converges absolutely and uniformly on the deleted interval (t_0, t_2). Therefore the function defined by the series $\sum_{h=1}^{\infty} y_h$ has a derivative at every point on the deleted interval (t_0, t_2), and this derivative is defined by the series $\sum_{h=1}^{\infty} y_h'$. This is true for every $\epsilon > 0$, and hence the function defined by the series (10) has a derivative at every point on the interval (t_0, t_2) except at the set E of measure zero.

It follows from the way in which the series (10) is derived that the function defined by this series satisfies the differential equation (1).

Conclusion. As there are two distinct solutions with the initial conditions $x(t_0) = 0$ (series (10) with $c = 0$), the point t_0 is a branch point of the solution of the differential equation (1). As t_0 can be any point in the interval for which the assumptions I, II, III are satisfied, it is clear that at every point on this interval there are two distinct solutions of the differential equation (1) which have the initial value $x(t_0) = 0$.

The case where the reciprocal of the right member of the differential equation (1) is analytic in x and t is well known.[3] It is evident that the conditions I, II, III do not depend upon analyticity and hence the method of this paper is more general.

HEIDELBERG COLLEGE.

[3] Ludwig Schlesinger, *Einführung in die Theorie der gewöhnlichen Differentialgleichungen auf funktiontheoretischer Grundlage*, p. 38.

FACTORIZATION IN PRINCIPAL IDEAL RINGS

By H. Serbin

Factorization theorems in a ring $K[x]$ of polynomials over a field K have been obtained by several writers [1, 3].[1] We consider this problem in the present paper for associative rings of a more general type, namely, rings (with unit element) in which every left ideal is a principal left ideal. We prove that two similar elements in such a ring which are factored in some way into a product of elements in $K[x]$ possess further factorizations which are essentially alike. The relation of two-sided ideals to left ideals is also considered for special rings. The latter leads to a generalization of a theorem due to N. Jacobson [2].

1. **Examples of rings which are principal left ideal rings.** Among such are:

(a) $K[x]$, a ring of polynomials $\alpha = \sum_{i=0}^{n} k_i x^i$ ($n = 0, 1, 2, \cdots, k_n \neq 0$) in one indeterminate x over a field K (non-commutative in general). The degree of α equals n. One assumes [3] that there is an associative and distributive (over addition) multiplication defined such that

$$\deg(\alpha \cdot \beta) = \deg \alpha + \deg \beta \tag{1}$$

for each pair α, β of elements of $K[x]$. Condition (1) implies

$$x\alpha = \bar{\alpha}x + \alpha'.$$

Cf. [2].

(b) R_n, the square matrix ring of degree n with elements from R, where R is itself a principal left ideal ring with unit element.

To show this, let I be a left ideal in R_n and consider the set of all vectors $(\alpha_1, \alpha_2, \cdots, \alpha_n)$, where $\alpha_k = a_{ik}$ ($k = 1, \cdots, n$), the matrix (a_{ik}) being in I. This set forms an R-left module M when addition is defined by

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) \pm (\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1 \pm \beta_1, \alpha_2 \pm \beta_2, \cdots, \alpha_n \pm \beta_n)$$

and

$$\rho(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\rho\alpha_1, \rho\alpha_2, \cdots, \rho\alpha_n), \qquad \rho \in R.$$

The first components of all the vectors in M thus form a left ideal in R generated by some element, say α, which must therefore be the first component of some vector A_1 in M. If $\alpha = 0$, then choose A_1 to be the null vector $(0, 0, \cdots, 0)$. Every vector of M is congruent (mod A_1) to a vector in which the first com-

Received January 31, 1938.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

ponent is zero. The vector A_2 is defined as one with first component zero and with second component a generator of the left ideal of all second components of vectors with zero first components. Continuing in this way, we obtain a sequence of vectors $A_1, A_2, \cdots, A_n$, some of which may be null vectors, such that

$$M \equiv 0 \pmod{A_1, A_2, \cdots, A_n}$$

so that $A_1, A_2, \cdots, A_n$ forms an R-basis. Then the matrix (a_{ij}) of degree n in which a_{ij} is the j-th component of A_i is the required generator of the left ideal I.

In the following, then, R will denote a principal left ideal ring (with unit element) which will be regarded upon occasion as an Abelian group under addition with left operator ring R itself [4]. The notation for the group theory and ideals as by van der Waerden [4] will be used here unless otherwise indicated. Thus the symbol (α) will mean the left ideal generated by $\alpha \in R$, the symbol $(\alpha)/(\beta)$ for $(\alpha) \supset (\beta)$ the quotient group in which (α), (β) are considered as normal subgroups of R.

In addition, a bar will be used to denote annihilators. More exactly, if $\beta \in R$, then the left ideal of all elements γ such that $\gamma\beta = 0$ will have a generator which will be denoted by $\bar{\beta}$.

2. Factorization by left ideals. If an element $\alpha \in R$ can be expressed as a product $\alpha_1\alpha_2 \cdots \alpha_n$, then we will say that α is *factored*, its *factors* being $\alpha_1, \alpha_2, \cdots, \alpha_n$. A factor α_j will be *trivial* if $(\alpha_j\alpha_{j+1} \cdots \alpha_n) = (\alpha_{j+1} \cdots \alpha_n)$.

LEMMA 1. *$\alpha = \beta^*\beta$ is equivalent to $(\beta) \supset (\alpha)$.*

LEMMA 2. *$\alpha = \beta_1^*\beta_2^* \cdots \beta_n^*$ is equivalent to $(\alpha) \subset (\beta_1) \subset (\beta_2) \subset \cdots \subset (\beta_{n-1})$, where $\beta_k = \beta_{k+1}^*\beta_{k+1}$ for $k = 0, 1, \cdots, n-1$, $\beta_0 = \alpha$, $\beta_n^* = \beta_{n-1}$.*

LEMMA 3. *If $\alpha = \beta^*\beta$, then*

$$(\beta)/(\alpha) \cong R/(\beta^*, \bar{\beta}) \qquad \textit{(operator-group isomorphism)}.$$

Since the first two lemmas are evident, only the third will be considered. The correspondence $\gamma \to \gamma\beta \pmod{(\alpha)}$ defines a homomorphism $R \sim (\beta)/(\alpha)$. The totality of elements γ of R going into the zero consists of all those such that $\gamma\beta \equiv 0 \pmod{(\alpha)}$ or

$$\gamma\beta \equiv 0 \pmod{(\beta^*\beta)},$$

that is,

$$\gamma\beta = \lambda\beta^*\beta \qquad \text{for some } \lambda,$$

whence

$$(\gamma - \lambda\beta^*)\beta = 0,$$

or

$$\gamma - \lambda\beta^* \equiv 0 \pmod{(\bar{\beta})}.$$

This gives finally

$$\gamma \equiv 0 \pmod{(\beta^*, \bar{\beta})}.$$

The result follows readily.

As usual, *similarity* is defined as follows:

DEFINITION. Two elements α and β in R are similar in R if

$$R/(\alpha) \cong R/(\beta).$$

Similarity is thus seen to satisfy the equivalence relations. A notable distinction between *equality* and *similarity*, however, lies in the fact that although α and β may not be similar in a ring R, they may be so in a ring $\bar{R}$ containing R. Such can occur in the case of *differential polynomials*.

Example. Let K be the field of rational functions of a variable t over the field of rational numbers. Consider the ring of differential polynomials $\sum_{i=0}^{n} a_i x^{n-i}$, where $x \equiv \dfrac{d}{dt}$ and $a_i \in K$. Multiplication is defined as is usual for linear differential operators. Primes denote differentiation with respect to t.

If P and Q are similar polynomials, then there are polynomials (see Theorem 3 below) P_1, Q_1 such that $PQ_1 = P_1Q$. If Q_1 is of higher degree than Q, write $Q_1 = SQ + T$, where the degree of T is lower than that of Q. Then $PT = T_1Q$, where $T_1 = -PS + P_1$. In particular, if $P = x - a$, $Q = x - b$, where $a, b \in K$, then T and T_1 are in K. But since $(x - a)T = Tx + T' - aT$, $T_1(x - b) = T_1x - T_1b$, it follows that $T = T_1$ and $T' = T(a - b)$. Conversely, if $T \in K$ satisfies this condition, then $x - a$, $x - b$ are similar (Theorem 3). Hence, if $a = 1$, $b = 0$, we conclude that $x - 1$, x are not similar. On the other hand, if K is enlarged to include e^t, these polynomials are similar.

THEOREM 1. *If α is similar to β and if $\alpha = \gamma_1^* \gamma_2^* \cdots \gamma_r^*$, $\beta = \delta_1^* \delta_2^* \cdots \delta_s^*$, then the factorization of each product can be continued so that*

$$\alpha = \alpha_1^* \alpha_2^* \cdots \alpha_n^*, \tag{2a}$$

$$\beta = \beta_1^* \beta_2^* \cdots \beta_n^*. \tag{2b}$$

Moreover, the new factors α^* in (2a) *can be paired with those β^* in* (2b) *so that, if α_j^* and β_k^* are corresponding factors,*

$$R/(\alpha_j^*, \bar{\alpha}_j) \cong R/(\beta_k^*, \bar{\beta}_k),$$

where $\bar{\alpha}_j$, $\bar{\beta}_k$ are annihilators of the products $\alpha_{j+1}^* \cdots \alpha_n^*$, $\beta_{k+1}^* \cdots \beta_n^*$, respectively.

Proof. Consider the group $S(\cong R/(\alpha) \cong R/(\beta))$. According to Lemma 2 we can construct a factor series of subgroups between (α) and R from the factorization given for α; similarly one for (β) and R. There are thus two factor

series between the null-group $0'$ of S and S. Apply Schreier's theorem [4] to these two series. They possess isomorphic refinements.

$$0' \subset (\alpha_1') \subset (\alpha_2') \subset (\alpha_3') \subset \cdots \subset (\alpha_{n-1}') = S, \tag{3a}$$

$$0' \subset (\beta_1') \subset (\beta_2') \subset (\beta_3') \subset \cdots \subset (\beta_{n-1}') = S. \tag{3b}$$

Let (α_j) be the left ideal containing all elements of R going into (α_j') under the homomorphism $R \sim R/(\alpha)$. Similarly with the β's. Interpret the series (3a) and (3b) in terms of (α_j) and (β_k).

$$(\alpha) \subset (\alpha_1) \subset (\alpha_2) \subset (\alpha_3) \subset \cdots \subset (\alpha_{n-1}) = R, \tag{4a}$$

$$(\beta) \subset (\beta_1) \subset (\beta_2) \subset (\beta_3) \subset \cdots \subset (\beta_{n-1}) = R. \tag{4b}$$

Since $(\alpha_j)/(\alpha_{j-1}) \cong (\alpha_j')/(\alpha_{j-1}')$ and similarly for the β's, Schreier's theorem implies that the factor series (4a) and (4b) are isomorphic. Hence, under a suitable pairing of quotient groups,

$$(\alpha_j)/(\alpha_{j-1}) \cong (\beta_k)/(\beta_{k-1}).$$

(Among such isomorphisms is $R/(\alpha_{n-1}) \cong R/(\beta_{n-1})$.) Let $\alpha_{j-1} = \alpha_j^*\alpha_j$, $\beta_{k-1} = \beta_k^*\beta_k$ $(j, k = 1, 2, \cdots, n-1)$, $\alpha_{n-1} = \alpha_n^*$, $\beta_{n-1} = \beta_n^*$, $\alpha_0 = \alpha$, $\beta_0 = \beta$, and apply Lemma 3. The preceding isomorphisms give

$$R/(\alpha_j^*, \bar{\alpha}_j) \cong R/(\beta_k^*, \bar{\beta}_k). \tag{5}$$

Before we proceed with a refinement of the theorem, it is convenient to state two lemmas.

LEMMA 4. *If* $(\alpha, \beta) = (\delta)$, *then* $\alpha = \alpha_1\delta$, $\beta = \beta_1\delta$ *and* $R = (\alpha_1, \beta_1, \bar{\delta})$.

Proof. For any particular $\rho \in R$ there exist elements $(')$, $('')$ in R such that

$$\rho\delta = (')\cdot\alpha + ('')\cdot\beta$$

by hypothesis. Lemma 1 gives $\alpha = \alpha_1\delta$, $\beta = \beta_1\delta$. Hence

$$\rho\delta = (')\alpha_1\delta + ('')\beta_1\delta,$$

and this gives $\rho - (')\alpha_1 - ('')\beta_1 \equiv 0 \pmod{(\bar{\delta})}$. Hence

$$\rho \equiv 0 \pmod{(\alpha_1, \beta_1, \bar{\delta})}.$$

LEMMA 5. *If* $\alpha = \beta^*\beta$, *then* $(\beta^*, \bar{\beta}) = R$ *is equivalent to* $(\alpha) = (\beta)$.

This follows from Lemma 3.

In the following, the symbol $(\gamma, -)$ to be used in connection with a definite factorization $\delta_1 \cdots \delta_{j-1}\gamma\delta_{j+1} \cdots \delta_n$ will denote the left ideal $(\gamma, \overline{\delta_{j+1} \cdots \delta_n})$. It is evidently independent of the factors preceding γ. The quotient groups $R/(\alpha_j^*, \bar{\alpha}_j)$ will be denoted therefore by $R/(\alpha_j^*, -)$.

We have the following

THEOREM 2. *The factorizations* (2a) *and* (2b) *of Theorem* 1 *may be replaced by*

(6a) $$\alpha = \lambda\lambda_1\lambda_2 \cdots \lambda_n ,$$

(6b) $$\beta = \mu\mu_1\mu_2 \cdots \mu_n ,$$

in which λ_j *is similar to* μ_k *for the pairs* j, k *as defined by Theorem* 1. *The factors* λ *and* μ *are trivial, i.e.,*

(7) $$(\alpha) = (\lambda_1\lambda_2 \cdots \lambda_n), \qquad (\beta) = (\beta_1\beta_2 \cdots \beta_n),$$

or what is equivalent

(8) $$R = (\lambda, -), \qquad R = (\mu, -).$$

Proof. The equivalence of (7) and (8) follows from Lemma 5. Start with the factorization (2a) $\alpha = \alpha_1\alpha_2 \cdots \alpha_n$, in which the asterisks have now been discarded. Let j be one of the numbers $1, 2, \cdots, n - 1$; let

(9) $$(\alpha_j, -) = (\alpha_j'),$$

and, consequently,

$$\alpha_j = \gamma\alpha_j', \qquad \overline{\alpha_{j+1} \cdots \alpha_n} = \delta\alpha_j'.$$

By Lemma 4,

(10) $$R = (\gamma, \delta, \bar{\alpha}_j').$$

Then

$$\alpha_j\alpha_{j+1} \cdots \alpha_n = \gamma\alpha_j'\alpha_{j+1} \cdots \alpha_n .$$

Consider $(\gamma, -)$. Evidently $(\gamma, -) \supset (\gamma, \delta, \bar{\alpha}_j') = R$ by (10). Hence $(\gamma, -) = R$. By Lemma 5, this gives

(11) $$(\alpha_j\alpha_{j+1} \cdots \alpha_n) = (\alpha_j'\alpha_{j+1} \cdots \alpha_n).$$

Now take $j > 1$. Then

(12) $$\begin{aligned} R/(\alpha_{j-1}, -) &\cong (\gamma\alpha_j'\alpha_{j+1} \cdots \alpha_n)/(\alpha_{j-1}\gamma\alpha_j'\alpha_{j+1} \cdots \alpha_n) \\ &\cong (\alpha_j'\alpha_{j+1} \cdots \alpha_n)/((\alpha_{j-1}\gamma)\alpha_j'\alpha_{j+1} \cdots \alpha_n) \cong R/((\alpha_{j-1}\gamma), -) \end{aligned}$$

if we apply Lemma 3, then (11), then Lemma 3 again. Therefore the factorization

(13) $$\alpha = \alpha_1\alpha_2 \cdots \alpha_{j-1}\alpha_j\alpha_{j+1} \cdots \alpha_n$$

is replaced by

(14) $$\alpha = \alpha_1\alpha_2 \cdots (\alpha_{j-1}\gamma)\alpha_j'\alpha_{j+1} \cdots \alpha_n$$

in which the $(j - 1)$-st factor is $\alpha_{j-1}\gamma$, the j-th being α_j'. What of the corresponding quotients $R/(*, -)$? (12) shows that these quotient groups correspond-

ing to the $(j-1)$-th factor are isomorphic before and after the replacement of α_j and regrouping. But $(\alpha_j', -) = (\alpha_j, -) = (\alpha_j')$ from (9). Hence

$$R/(\alpha_j, -) \cong R/(\alpha_j', -) \cong R/(\alpha_j').$$

The final result follows immediately. Start with $j = n - 1$ in the factorization (13) and perform the replacement (13) $\rightarrow$ (14). Repeat for $j = n - 2$, etc. Finally we arrive at the factorizations of (6a) and (6b). Equation (11) gives (7) and (8) of the theorem.

3. **Similarity**. The following theorem gives a well-known interpretation of similarity in terms of operations and elements of the ring R itself.

THEOREM 3. *If R has a unit element* 1, *then all of the following statements are equivalent*:

(a) $R/(\alpha) \cong R/(\beta)$.

(b) *There is an element θ in R such that* (i) $(\alpha\theta) = [\beta, \theta]$, (ii) $(\beta, \theta) = R$, *and* (iii) *every left annihilator of θ is in (α).*

(c) *There is an element φ in R such that* (i) $(\beta\varphi) = [\alpha, \varphi]$, (ii) $(\alpha, \varphi) = R$, *and* (iii) *every left annihilator of φ is in (β).*

Proof. (b) and (c) can be obtained from one another by interchange of α, β and φ, θ.

Assume (b). Then (b), (iii) implies that $(\alpha, \bar{\theta}) = (\alpha)$. Then

$$R/(\alpha) \cong (\theta)/(\alpha\theta) \cong (\theta)/[\beta, \theta] \cong (\beta, \theta)/(\beta) \cong R/(\beta). \tag{15}$$

The isomorphism $R/(\alpha) \cong R/(\beta)$ is seen to be given by

$$\gamma \pmod{(\alpha)} \rightleftarrows \gamma\theta \pmod{(\beta)}. \tag{16}$$

Assume (a). Then $1 \pmod{(\alpha)} \rightleftarrows \theta \pmod{(\beta)}$ for some θ in R. Since $\gamma = \gamma \cdot 1$,

$$\gamma \pmod{(\alpha)} \rightleftarrows \gamma\theta \pmod{(\beta)}, \tag{17}$$

and, in particular, $\alpha \pmod{(\alpha)} \rightleftarrows \alpha\theta \pmod{(\beta)}$ so that

$$\alpha\theta \equiv 0 \pmod{(\beta)}. \tag{18}$$

The isomorphism implies that there is a $\varphi \in R$ such that

$$\varphi\theta \equiv 1 \pmod{(\beta)},$$

$$(\theta, \beta) = R. \tag{19}$$

But the correspondence (17) is such that if $\gamma\theta \equiv 0 \pmod{(\beta)}$, then $\gamma \equiv 0 \pmod{(\alpha)}$. This gives (iii) of (b). Finally, that (i) follows can be shown by considering a sequence of isomorphisms as in (15). We see that $(\theta)/(\alpha\theta) \cong (\theta)/[\beta, \theta]$ under the correspondence $\gamma\theta \pmod{(\alpha\theta)} \rightleftarrows \gamma\theta \pmod{[\beta, \theta]}$, and this gives the desired result.

If $R/(\alpha)$ were homomorphic to a subgroup of $R/(\beta)$, we would have merely (18). Hence we have

THEOREM 3a. *The set H of homomorphisms of $R/(\alpha)$ into itself forms a ring which is isomorphic to the ring* (mod (α)) *of elements θ of R such that $\alpha\theta \equiv 0$* (mod (α)).

THEOREM 3b. *If (α) is divisorless, then H is a field.*

4. **Two-sided ideals.** Let (α) be a two-sided ideal in R. Consider the quotient ring $R//(\alpha)$. There is a (1-1) correspondence between left (right) ideals of R which contain (α) and all left (right) ideals in $R//(\alpha)$. The ring $R//(\alpha)$ is subjected to the following

Condition. $R//(\alpha)$ satisfies a descending chain condition for left ideals. This is satisfied if R is a principal left ideal and principal right ideal ring which is a domain of integrity.

THEOREM 4. *Let R be a domain of integrity, (α) a two-sided ideal such that the above condition is satisfied. If (α) is not divisible by any two-sided ideal other than (α) and R, then α can be written as a product $\alpha = \alpha_1\alpha_2 \cdots \alpha_n$, where each α_k can be factored no further. All the α_k's so obtained are similar.*

This is a generalization of a theorem of N. Jacobson [2, p. 199].

Proof. $R//(\alpha)$ is a simple ring satisfying descending and ascending chain conditions for left ideals. Then

$$R^* \cong R//(\alpha) \cong l_1 + l_2 + l_3 + \cdots + l_n,$$

where the l's are minimal left ideals which, considered as groups with left operator ring $R//(\alpha)$, are operator-isomorphic to each other. To the series of composition in R^*

$$0 < l_1 < l_1 + l_2 < \cdots < l_1 + l_2 + \cdots + l_n = R^*$$

there corresponds a series of composition in R

$$(\alpha) < (\alpha_1') < (\alpha_2') < \cdots < (\alpha_n') = R, \tag{20}$$

where (α_k') is the left ideal in R consisting of all elements going into $l_1 + l_2 + \cdots + l_k$ under the ring homomorphism $\rho \to r$ defined by $R \to R^*$. Hence

$$(\alpha_k')/(\alpha_{k-1}') \cong (l_1 + l_2 + \cdots + l_k)/(l_1 + l_2 + \cdots + l_{k-1}) \cong l_k, \tag{21}$$

the isomorphisms being for *addition*. Now R^* is a group under addition with left operator ring R^*. It can also be defined as a group under addition with left operator ring R. For this, let ρ be in R, a in R^*. Then define

$$\rho \cdot a = r \cdot a,$$

where $\rho \to r$ under the homomorphism $R \to R^*$. Also, the isomorphism of the minimal left ideals gives rise to an isomorphism between the subgroups $R \cdot l_k$. For if $a_i \in l_i$, $a_k \in l_k$ such that $a_i \rightleftarrows a_k$ in the isomorphism $l_i \cong l_k$ (preserved

under the operator ring R^*), then $\rho a_i \rightleftarrows \rho a_k$. Since (α_k') is the totality of elements of R going into $l_1 + \cdots + l_k = R^*(l_1 + \cdots + l_k) = R(l_1 + \cdots + l_k)$, then the isomorphism of (21) is also an operator-group isomorphism. The result follows by an application of Lemmas 2 and 3.

We also obtain

THEOREM 5. *Under the conditions of Theorem 4, α is the left (right) least common multiple of $\alpha_1, \alpha_2, \cdots, \alpha_n$. The α's are left (right) relatively prime in the sense $(\alpha_j, \alpha_k)_l = R$ $((\alpha_j, \alpha_k)_r = R)$.*

BIBLIOGRAPHY

1. H. BLUMBERG, *Über algebraische Eigenschaften von linearen Differentialausdrücken*, Dissertation, Göttingen, 1912.
2. N. JACOBSON, *Non-commutative polynomials and cyclic algebras*, Annals of Mathematics, vol. 35(1934), pp. 197–208.
3. O. ORE, *Theory of non-commutative polynomials*, Annals of Mathematics, vol. 34(1933), pp. 480–508. See also Journal für Mathematik, vols. 167, 168.
4. B. L. VAN DER WAERDEN, *Moderne Algebra*, vols. I and II, Berlin, 1930.

INSTITUTE FOR ADVANCED STUDY.

NIL-RINGS WITH MINIMAL CONDITION FOR ADMISSIBLE LEFT IDEALS

By Charles Hopkins

The main theorem of this article is stated in §1 and proved in §2. Possibly the corollaries of this theorem are of more interest than the theorem itself. Let $\mathfrak{O}$ be any ring with minimal conditions for left ideals. From our main theorem it follows that (1) the radical of $\mathfrak{O}$ is nilpotent; (2) the ring $\mathfrak{O}$ is semi-primary (or semi-simple); (3) any subring of $\mathfrak{O}$ containing only nilpotent elements is itself nilpotent. This third corollary is a conjecture of Köthe, which Levitzki[1] proved by assuming both the minimal and maximal condition for right ideals of $\mathfrak{O}$.

1. **Definitions and assumptions**. Let $\mathfrak{N}$ be a nil-ring—i.e., a ring in which every element is nilpotent—and let Ω denote a set of operators for $\mathfrak{N}$, each element of Ω being a left-hand operator for $\mathfrak{N}$. We shall assume that (1) $\mathfrak{N}$ is not the null-ring and that (2) the set Ω contains all the elements of $\mathfrak{N}$. Thus Ω will contain as right-hand operators the elements of $\mathfrak{N}$ (and possibly elements not belonging to $\mathfrak{N}$). We assume the following postulates:

(α_0) $\xi(u + v) = \xi u + \xi v$ for all $\xi \in \Omega$ and u, v in $\mathfrak{N}$;

(α_1) $(\xi\eta)u = \xi(\eta u)$, provided that $\xi\eta$ exists in Ω;

(α_2) $(\xi + \eta)u = \xi u + \eta u$, if $\xi + \eta$ is defined in Ω.

For those elements of Ω which are right-hand operators for $\mathfrak{N}$ we assume the analogues of (α_0)-(α_2) above; e.g., (α_1') asserts that $u(\xi\eta) = (u\xi)\eta$, provided that the product $\xi\eta$ exists in Ω and is a right-hand operator.

If an element θ of Ω is *not* a right-hand operator for $\mathfrak{N}$, we shall need the additional postulate:

(α_3) $\theta(uv) = u(\theta v)$.

At this point we mention three useful relations which are consequences of (2), (α_1), and (α_1') above:

(β) $\xi(uv) = (\xi u)v$;

(γ) $(v\xi)u = v(\xi u)$;

(δ) $(vu)\xi = v(u\xi)$.

We derive (β) from (α_1), and (γ) and (δ) from (α_1'), by regarding the element u of $\mathfrak{N}$ as an operator (see (2) above). Obviously (β) holds for all ξ in Ω, while (γ) and (δ) are valid only when ξ is a right-hand operator. In connection with (α_3) we point out that if ξ is a right-hand operator we do not deny (α_3)—we merely do not assume it.

Received February 28, 1938. I am indebted to members of the Mathematics Seminar of Tulane University for useful suggestions regarding the phraseology of the proofs.

[1] Math. Ann., vol. 105(1931), pp. 620–627.

By the term *admissible left ideal* of $\mathfrak{R}$ (a.l.i.) we shall mean a left ideal of $\mathfrak{R}$ which admits all the elements of Ω as left-hand operators.

Fundamental Hypothesis. We assume that the descending chain condition (minimal condition) holds for admissible left ideals of $\mathfrak{R}$.

PRINCIPAL THEOREM. *The ring $\mathfrak{R}$ is nilpotent.*

2. **Proof of the Principal Theorem.** Our proof will depend upon the following auxiliary theorems:

2.1. *The ring $\mathfrak{R}^{2^m}$ $(m = 0, 1, 2, \cdots)$ admits the elements of Ω as operators and satisfies the same assumptions and postulates (relative to Ω) as does $\mathfrak{R}$ itself (with the possible exception of* (1)).

2.2. *If $\mathfrak{R}^{2^m} \supset 0$, then $\mathfrak{R}^{2^m} \supset (\mathfrak{R}^{2^m})^2$.*

Our main theorem will be a direct consequence of these two lemmas; for 2.1 implies that each $\mathfrak{R}_m$ $(= \mathfrak{R}^{2^m})$ is an admissible left ideal of $\mathfrak{R}$, and since the minimal condition holds for admissible left ideals of $\mathfrak{R}$, we know that in the chain $\mathfrak{R} = \mathfrak{R}_0 \supseteqq \mathfrak{R}_1 \supseteqq \cdots$ we must have the equality sign after a finite number of terms. But from 2.2 we know that $\mathfrak{R}_m$ can equal $\mathfrak{R}_{m+1}$ only if $\mathfrak{R}_m = 0$. Hence $\mathfrak{R}$ is nilpotent, and its exponent does not exceed 2^m.

Proof of 2.1. Evidently $\mathfrak{R}^n$ is a two-sided ideal of $\mathfrak{R}$ for all positive integral values of n. Let ξ and $x_1x_2 \cdots x_n$ be any elements of Ω and $\mathfrak{R}^n$, respectively. From (β) of §1 we have

$$\xi(x_1x_2 \cdots x_n) = (\xi x_1)(x_2 \cdots x_n) \in \mathfrak{R}^n.$$

If ξ is a right-hand operator of $\mathfrak{R}$, then, using (δ) of §1, we obtain

$$(x_1x_2 \cdots x_n)\xi = (x_1 \cdots x_{n-1})(x_n\xi) \in \mathfrak{R}^n.$$

Hence a left-hand (right-hand) operator for $\mathfrak{R}$ is a left-hand (right-hand) operator for $\mathfrak{R}^n$. In particular, $\mathfrak{R}^n$ is an admissible left ideal of $\mathfrak{R}$. Since $\mathfrak{R}^n$ is a subring of $\mathfrak{R}$, all the postulates which are satisfied by the elements of $\mathfrak{R}$ are automatically satisfied by the elements of $\mathfrak{R}^n$.

That the minimal condition holds for admissible left ideals of $\mathfrak{R}^n$ follows from the fact that every a.l.i. of $\mathfrak{R}^n$ is an a.l.i. of $\mathfrak{R}$ (see (2) of §1).

Proof of 2.2. We assume that $\mathfrak{R}_m$ $(= \mathfrak{R}^{2^m})$ is not the null-ring. We wish to show that $\mathfrak{R}_m \supset \mathfrak{R}_m^2$. Since $\mathfrak{R}_m$ and $\mathfrak{R}$ satisfy the same postulates relative to Ω (see 2.1), we may drop the subscript and prove 2.2 for $\mathfrak{R}$ alone. We shall need the following lemmas, the proofs of which we shall postpone.

2.21. *If $\mathfrak{l}$ is a minimal non-zero admissible left ideal of $\mathfrak{R}$ (m.a.l.i.), then $\mathfrak{R}\mathfrak{l} = 0$.*

DEFINITION. The *right annihilator* of $\mathfrak{R}$ is the set of all elements u in $\mathfrak{R}$ for which $\mathfrak{R}u = 0$.

2.22. *The right annihilator $\mathfrak{N}$ of $\mathfrak{R}$ is* (a) *a two-sided ideal of $\mathfrak{R}$ and* (b) *an admissible left ideal of $\mathfrak{R}$.*

We know that $\mathfrak{N} \supset 0$, since $\mathfrak{N}$ must contain every non-zero m.a.l.i. If $\mathfrak{N} = \mathfrak{R}$, then $\mathfrak{R}^2 = \mathfrak{R}\mathfrak{N} = 0 \subset \mathfrak{R}$, and our proof is ended. So we assume $\mathfrak{N} \subset \mathfrak{R}$. Since $\mathfrak{N}$ is a two-sided ideal of $\mathfrak{R}$, there exists the quotient-ring $\mathfrak{R}' = \mathfrak{R}/\mathfrak{N}$. Let x'

denote the element of $\mathfrak{N}'$ that corresponds to the element x of $\mathfrak{N}$ in the homomorphism $\mathfrak{N} \sim \mathfrak{N}'$. We shall find it convenient to denote the element x' by the class $x + \mathfrak{R}$ of elements in $\mathfrak{N}$; in particular, the null-element $0'$ of $\mathfrak{N}'$ will be the class $\mathfrak{R}$.

Since $\mathfrak{R}$ is an a.l.i. of $\mathfrak{N}$, we may define the product $\xi x'$, for all $\xi \in \Omega$, by the equation $\xi x' = \xi x + \mathfrak{R} = (\xi x)'$. If ξ is a right-hand operator of $\mathfrak{N}$, we have the corresponding definition[2] $x'\xi = (x\xi)'$.

Thus a left-hand (right-hand) operator for $\mathfrak{N}$ is a left-hand (right-hand) operator for $\mathfrak{N}'$. Every element of $\mathfrak{N}'$ is nilpotent; for if ρ is the exponent of x, then $x'^{\rho} = (x + \mathfrak{R})^{\rho} = x^{\rho} + \mathfrak{R} = \mathfrak{R} = 0'$. One can easily verify the fact that $\mathfrak{N}'$ satisfies the same postulates relative to Ω as does $\mathfrak{N}$ itself. It is known that the minimal condition for a.l. ideals of a ring $\mathfrak{O}$ implies the minimal condition for a.l. ideals of any quotient-ring $\mathfrak{O}' = \mathfrak{O}/\mathfrak{a}$ (it being assumed, naturally, that $\mathfrak{O}$ and $\mathfrak{O}'$ have the same domain of operators). Hence we may deduce, from 2.21 and 2.22, the existence of a right annihilator $\mathfrak{R}' \supset 0'$ (for the ring $\mathfrak{N}'$) which is an a.l.i. of $\mathfrak{N}'$.

Let $\mathfrak{R}_1$ denote the set of all elements in $\mathfrak{N}$ which correspond to elements of $\mathfrak{R}'$ in the homomorphism $\mathfrak{N} \sim \mathfrak{N}'$. Since $\mathfrak{R}' \supset 0'$, we have $\mathfrak{R}_1 \supset \mathfrak{R}$.

Now $\mathfrak{N}\mathfrak{R}_1$ corresponds to $\mathfrak{N}'\mathfrak{R}'$ $(= 0')$ in the homomorphism $\mathfrak{N} \sim \mathfrak{N}'$; hence $\mathfrak{N}\mathfrak{R}_1 \subseteq \mathfrak{R}$. Therefore $\mathfrak{N}^2\mathfrak{R}_1 \subseteq \mathfrak{N}\mathfrak{R} = 0$. But $\mathfrak{N}\mathfrak{R}_1 \neq 0$; otherwise, we should have $\mathfrak{R}_1 \subseteq \mathfrak{R}$, from the definition of the right annihilator $\mathfrak{R}$. Since $\mathfrak{N}\mathfrak{R}_1 \neq 0$ and $\mathfrak{N}^2\mathfrak{R}_1 = 0$, we must obviously have the inequality $\mathfrak{N} \supset \mathfrak{N}^2$.

Proof of 2.21. Let $\mathfrak{l}$ be any non-zero m.a.l.i. of $\mathfrak{N}$ and let u denote any non-zero element of $\mathfrak{l}$. Now $\mathfrak{N}u$ is an a.l.i. of $\mathfrak{N}$, and since $\mathfrak{l}$ is minimal, we must have either $\mathfrak{N}u = \mathfrak{l}$ or $\mathfrak{N}u = 0$. We exclude the first alternative by proving that $\mathfrak{N}u$, which is contained in $\mathfrak{l}$, cannot contain u. Suppose that $u = au$ for some a in $\mathfrak{N}$, and let ρ denote the exponent of a. Then we should have $u = au = a^2u = \cdots = a^{\rho}u = 0$, contrary to our assumption $u \neq 0$. Since u represents *any* non-zero element of $\mathfrak{l}$, we conclude that $\mathfrak{N}\mathfrak{l} = 0$.

Proof of 2.22. We wish to show that the right annihilator $\mathfrak{R}$ of $\mathfrak{N}$ is (a) a two-sided ideal of $\mathfrak{N}$ and (b) an admissible left ideal of $\mathfrak{N}$. Since Ω contains all the elements of $\mathfrak{N}$, it will be sufficient to show that (i) a left-hand operator of $\mathfrak{N}$ is a left-hand operator of $\mathfrak{R}$; (ii) a right-hand operator of $\mathfrak{N}$ is a right-hand operator of $\mathfrak{R}$.

Let ξ, x, and u be any elements of Ω, $\mathfrak{N}$, and $\mathfrak{R}$, respectively.

(i) If ξ is not a right-hand operator of $\mathfrak{N}$, then, using postulate (α_3) of §1, we obtain $x(\xi u) = \xi(xu) = \xi \cdot 0 = 0$. That is, ξu annihilates every element of $\mathfrak{N}$ on the right, and is therefore contained in $\mathfrak{R}$.

If ξ is a right-hand operator of $\mathfrak{N}$, then, from (γ) of §1, we have $x(\xi u) = (x\xi)u \in \mathfrak{N}u = 0$.

(ii) If ξ is a right-hand operator, then, from (δ) of §1, we have $x(u\xi) = (xu)\xi = 0 \cdot \xi = 0$; that is, $u\xi \in \mathfrak{R}$.

[2] In the proof of 2.22 we shall show that $\mathfrak{R}$ admits every right-hand operator of $\mathfrak{N}$.

3. **Consequences of the Principal Theorem**. Let $\mathfrak{O}$ be any ring with minimal condition for left ideals, and let $\mathfrak{R}$ denote the radical of $\mathfrak{O}$. In what follows $\mathfrak{O}$ assumes the rôle of Ω above.

3.1. *The radical of $\mathfrak{O}$ is nilpotent.*

This result is obviously nothing more than a restatement of our main theorem for the case $\Omega = \mathfrak{O}$.

3.2. *A ring $\mathfrak{O}$ with radical and with minimal condition for left ideals is semi-primary.*

By definition, a semi-primary ring is one for which the quotient-ring with respect to the radical is semi-simple. Let $\mathfrak{O}'$ denote the quotient-ring $\mathfrak{O}/\mathfrak{R}$. Since the minimal condition for left ideals holds in $\mathfrak{O}$, it will hold in $\mathfrak{O}'$. Let us suppose that $\mathfrak{O}'$ contains a radical $\mathfrak{R}' \supset 0'$. From 3.1 we know that $\mathfrak{R}$ and $\mathfrak{R}'$ are both nilpotent. Let ρ and ρ' denote the exponents of $\mathfrak{R}$ and $\mathfrak{R}'$, respectively. Now the set $\mathfrak{R}_1$ of all elements in $\mathfrak{O}$ which correspond to elements of $\mathfrak{R}'$ is a two-sided ideal of $\mathfrak{O}$. Since $\mathfrak{R}_1^{\rho'\rho} \subseteq \mathfrak{R}^\rho = 0$, we see that $\mathfrak{R}_1$ must be nilpotent. But the radical $\mathfrak{R}$ of $\mathfrak{O}$ must contain every nilpotent two-sided ideal; we must have, accordingly, $\mathfrak{R} \supseteq \mathfrak{R}_1$. But if $\mathfrak{R}' \supset 0'$, then $\mathfrak{R}_1 \supset \mathfrak{R}$. It follows, therefore, that $\mathfrak{O}/\mathfrak{R}$ is a ring without radical and with minimal condition for left ideals—i.e., $\mathfrak{O}$ is semi-primary.

3.3. *In a ring $\mathfrak{O}$ with minimal condition for left ideals any subring $\mathfrak{S}$ containing only nilpotent elements is itself nilpotent.*

(i) If $\mathfrak{O}$ is semi-simple, then it is well known that the maximal condition holds for left ideals of $\mathfrak{O}$. In this case, therefore, Levitzki's proof is available.[3]

(ii) If $\mathfrak{S}$ is contained in the radical $\mathfrak{R}$ of $\mathfrak{O}$, then $\mathfrak{S}$ is necessarily nilpotent (see 3.1).

(iii) We assume $\mathfrak{R} \supset 0$ and $\mathfrak{S} \not\subseteq \mathfrak{R}$.

Let $\mathfrak{T}$ denote the set of all elements $x_s + y_r$, $x_s \in \mathfrak{S}$, $y_r \in \mathfrak{R}$. Since $\mathfrak{R}$ is a two-sided ideal of $\mathfrak{O}$, it is easy to see that $\mathfrak{T}$ is a ring containing both $\mathfrak{R}$ and $\mathfrak{S}$ as subrings.

Let $\mathfrak{T}'$ denote the subring of $\mathfrak{O}'$ which corresponds to $\mathfrak{T}$ in the homomorphism $\mathfrak{O} \sim \mathfrak{O}' = \mathfrak{O}/\mathfrak{R}$. Let σ denote the exponent of x_s. Then $(x_s + y_r)^\sigma \in \mathfrak{R}$, since $\mathfrak{R}$ is a two-sided ideal of $\mathfrak{O}$. It follows, then, that every element of $\mathfrak{T}'$ is nilpotent; since $\mathfrak{O}'$ is semi-simple (see 3.2), we conclude from (i) above that $\mathfrak{T}'$ is nilpotent. Let ρ' and ρ denote the exponents of $\mathfrak{T}'$ and $\mathfrak{R}$, respectively. Then $\mathfrak{T}^{\rho'\rho} \subseteq \mathfrak{R}^\rho = 0$. That is, $\mathfrak{T}$ is nilpotent. And since $\mathfrak{S}$ is contained in $\mathfrak{T}$, it follows immediately that $\mathfrak{S}$ is also nilpotent.

TULANE UNIVERSITY.

[3] See footnote 1.

INFINITE SYSTEMS OF LINEAR EQUATIONS AND EXPANSIONS OF ANALYTIC FUNCTIONS

By P. W. Ketchum

Introduction. In the present article we emphasize a mutual relationship existing between the theory of the solution of infinite systems of linear equations in an infinity of unknowns and the theory of the expansion of an analytic function in a series of analytic functions. In the first two sections we show how known expansion theorems can be used to give theorems on the solution of infinite systems of equations. In particular, an expansion theorem of Birkhoff[1] yields a theory similar to that of von Koch[2] on normal determinants.

In the remainder of the paper we apply known theorems, or suitable modifications of known theorems, on infinite systems of linear equations to new situations so as to obtain generalizations of certain types of expansion theorems. These expansions are similar in character to those of Pincherle,[3] who was the first to show that any function $f(x)$ analytic at $x = 0$ can be expanded in a series of the form

$$f(x) = \sum_{m=0}^{\infty} c_m x^m G_m(x), \qquad G_m(0) = 1, \tag{1}$$

provided the functions $G_m(x)$ are analytic and uniformly bounded in some neighborhood of $x = 0$.

We have adopted in this part of the paper a point of view initiated by I. M. Sheffer,[4] according to which a known expansion theorem is generalized by replacing the functions, in terms of which we are expanding, by linear combinations of those same functions, the coefficients of these linear combinations being restricted by appropriate conditions so that the resulting sums will be "close to" the original functions.

In the expansions of Pincherle, the n-th term has a zero of order n at the origin. Our expansions differ from those of Pincherle principally in that the n-th term has n zeros not necessarily coincident.

Received April 8, 1938.

[1] Comptes Rendus, vol. 164(1917), pp. 942–945.

[2] See F. Riesz, *Les Systèmes d'Équations Linéaires à une Infinité d'Inconnues*, Paris, 1913, Chapter II.

[3] Memorie della Accademia delle Scienze dell' Istituto di Bologna, (4), vol. 3(1881), pp. 151 ff.

[4] American Journal of Mathematics, vol. 57(1935), pp. 587–614.

1. Let $f(x)$ and $F_m(x)$ be given functions of the complex variable x,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \tag{2}$$

$$F_m(x) = \sum_{n=0}^{\infty} b_{m,n} x^n \qquad (m = 0, 1, \cdots), \tag{3}$$

where the radii of convergence of these series are not zero. If it is possible to have the uniformly convergent expansion

$$f(x) = \sum_{m=0}^{\infty} c_m F_m(x), \tag{4}$$

then, according to the Weierstrass double series theorem, the coefficients of the various powers of x, obtained by substituting the expansions of $f(x)$ and $F_m(x)$, must be equal on both sides of the equation. In other words, the equations

$$\sum_{m=0}^{\infty} c_m b_{m,n} = a_n \qquad (n = 0, 1, \cdots) \tag{5}$$

are valid. When the a's and b's are regarded as given, these form an infinite system of linear equations in the infinity of unknown c's. Under our hypothesis that the expansion (4) holds uniformly in some neighborhood of $x = 0$, the system (5) has a solution. Moreover, information concerning the asymptotic behavior of the solution is given implicitly, but none the less definitely, by the condition that the inequality

$$\limsup_{m\to\infty} | c_m F_m(x) |^{1/m} \leqq 1 \tag{6}$$

hold for x in the region of convergence of (4).

In important classes of cases we may translate this implicit relation into an explicit condition on the c's. For example, suppose that the functions $F_m(x)$ are such that

$$\lim | F_m(x) |^{-1/m} = \phi(x) \tag{7}$$

for x in a region S, about the origin, in which $f(x)$ and every $F_m(x)$ are analytic. It is understood that $\phi(x)$ may have infinity as a value. Suppose further that (4) holds uniformly in S. Then

$$\limsup | c_m |^{1/m} \leqq L, \tag{8}$$

where L is the lower bound (finite or infinite) of $\phi(x)$ in S. We can thus state

THEOREM I. *The infinite system of equations* (5), *with the c's as unknowns, has a solution provided* (4) *holds uniformly in some neighborhood of the origin. This solution satisfies* (8) *provided* (4) *is satisfied uniformly in a region S about the origin in which* $f(x)$ *and* $F_m(x)$ *are analytic and* (7) *holds.*

If the equations (5) are such that the functions $f(x)$ and $F_m(x)$ exist, then

these functions will be said to be *associated* functions, that is, associated with the system (5). If the associated functions do not exist, it may be possible to divide the equations by appropriate constants so as to get a new system for which they do exist. In particular, if all but a finite number of them exist, then one can always find multipliers such that the new system will have a complete set of associated functions.

As an example of an application of Theorem I, consider the particular system of equations (5) in which $a_n = 1$, $b_{m,0} = 1$, $b_{0,n} = 0$, $n > 0$, and $b_{m,n}$ for $m > 0$, $n > 0$ is the sum of all possible products of $\frac{1}{2}, \frac{1}{3}, \cdots, 1/m$ taken n at a time (allowing repetitions). Since $b_{m,1}$ does not approach zero, the infinite determinant $|b_{m,n}|$ is not normal, nor can it be made normal by introducing multipliers for the equations. The associated functions are

$$f(x) = (1 - x)^{-1},$$

$$F_0(x) = 1, \qquad F_m(x) = \left[\left(1 - \frac{x}{2}\right)\left(1 - \frac{x}{3}\right)\cdots\left(1 - \frac{x}{m}\right)\right]^{-1}.$$

The expansion (4) will then be an expansion in Stirling's series, which will be valid and uniformly convergent on any closed set of points in the half plane $R(x) < 1$. Moreover, $\phi(x) \equiv 1$. Hence this particular infinite system of equations has a solution satisfying the conditions

$$\limsup |c_n|^{1/n} \leqq 1.$$

2. As another application of Theorem I, we use the following expansion theorem of Birkhoff:

If the functions $f(x)$ and $F_m(x)$ are analytic for $|x| \leqq a$, and if

$$\sum_{m=0}^{\infty} |F_m(x) - x^m| \, a^{-m-1} \tag{9}$$

converges uniformly for $|x| = a$ to a value less than a^{-1}, then (4) holds uniformly for the same circular region $|x| \leqq a$.

Hence we have the following corollary to Theorem I:

THEOREM II. *If the functions associated with the given equations* (5) *are such that* (9) *converges uniformly to a value less than a^{-1} for a region $|x| \leqq a$ in which $f(x)$ and $F_m(x)$ are analytic, then* (5) *has a solution satisfying* (6) *in that region. Furthermore, if* (7) *is satisfied for $|x| \leqq a$, then the solution will satisfy* (8).

The special case of Theorem II where $a = 1$ is of interest because of its similarity to the classical theorem of von Koch that (5) has a bounded solution provided the a's are bounded and the determinant of the equations is normal and different from zero. The condition that (9) converge for $|x| = a = 1$ resembles the condition that the determinant of the equations be normal. Thus, for $a = 1$, Theorem II requires that the series (9) for $x = 1$,

$$\sum_{m=0}^{\infty} \left| \sum_{n=0}^{\infty} (b_{m,n} - \delta_{m,n}) \right|,$$

converge, $\delta_{m,n}$ being Kronecker's symbol; while the condition that the determinant be normal is that the series

$$\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} | b_{m,n} - \delta_{m,n} |\right)$$

converge.

Theorem II possesses certain advantages over von Koch's theorem in that the condition imposed on the a's is less restrictive, and it may be easier to show that (9) has a value less than a^{-1} than to show that the determinant of the coefficients does not vanish.

3. We now consider the problem inverse to that just discussed. Suppose that the equations (5), in which the c's are regarded as unknowns, have a solution. Then (4) will be a formal expansion of $f(x)$. Suppose, moreover, that the solution of (5) is known to satisfy the condition

$$\limsup | c_m |^{1/m} \leqq L \tag{10}$$

for some number L, and that $f(x)$ and $F_m(x)$ are analytic in a region S about the origin. Then the expansion will be valid uniformly in any closed region about the origin, in S, where

$$\limsup | F_m(x) |^{1/m} \leqq K < L^{-1}. \tag{11}$$

We can now state

THEOREM III. *Let the equations* (5) *be such that there is a solution satisfying* (10). *Suppose further that the associated functions exist and are analytic in a region S, about the origin. Then* (4) *will be a valid uniformly convergent expansion in any closed region, about the origin and in S, where* (11) *holds.*

As an application of Theorem III, we may suppose that the determinant $| b_{m,n} |$ is normal and different from zero. If, then, the a's are bounded, there will be a solution with the c's also bounded. In this case $L \leqq 1$.

4. From this point on, we propose to restrict attention to a special class of equations (5) which are, in fact, merely recurrence relations. In such a case there is, of course, no question as to the existence of a solution. The problem of determining the asymptotic behavior of the solution is, however, of as much interest for this case as for the more general class of equations. The equations are as follows:

$$\begin{aligned} c_0 &= a_0\,, \\ \beta_{0,1}c_0 + c_1 &= a_1\,, \\ &\cdots\cdots\cdots\cdots \\ \beta_{0,i}c_0 + \cdots + \beta_{i-1,1}c_{i-1} + c_i &= a_i \\ &\cdots\cdots\cdots\cdots\,. \end{aligned} \tag{12}$$

They may be obtained by writing

$$G_m(x) = \sum_{n=0}^{\infty} \beta_{m,n} x^n, \qquad \beta_{m,0} = 1,$$

and equating coefficients of corresponding powers of x in Pincherle's expansion (1). Corresponding to Pincherle's theorem, stated in the introduction, we have the following

THEOREM IV. *If in the equations* (12)

$$\beta_{n,m} = O(\rho^{-m}) \qquad \textit{uniformly in } n, \tag{13}$$

$$a_m = O(r^{-m}), \tag{14}$$

then there will exist a γ *such that*

$$c_m = O(\gamma^{-m}). \tag{15}$$

This is a special case of the following theorem, which corresponds to an expansion theorem of Takenaka.[5]

THEOREM V. *If in the equations* (12)

$$\beta_{n,m} = O(\rho^{-m}), \qquad i.e., \ |\beta_{n,m}| \leqq M_n \rho^{-m}, \tag{16}$$

$$a_m = O(r^{-m}), \tag{17}$$

then, letting ϵ *be any positive number and* $M = \limsup M_n$, *we have*

$$c_m = O(\max[\rho^{-m}(M + 1 + \epsilon)^m, r^{-m}]). \tag{18}$$

Proof. Consider the sequence of systems of equations obtained from the system (12) by modifying the right members so that $a_i = \delta_{i,j}$ for the j-th member of the sequence. Let $\{c_i^{(j)}\}$ be the solution of the j-th such system. Evidently $c_i^{(j)} = 0$ for $i < j$. The solution of (12) will be given by the sum

$$c_i = \sum_{j=0}^{i} c_i^{(j)} a_j. \tag{19}$$

We first estimate the numbers $c_i^{(j)}$. We have

$$c_{j+\nu}^{(j)} = (-1)^\nu \begin{vmatrix} \beta_{j,1} & 1 & 0 & \cdots & 0 & 0 \\ \beta_{j,2} & \beta_{j+1,1} & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{j,\nu-1} & \beta_{j+1,\nu-2} & \beta_{j+2,\nu-3} & \cdots & \beta_{j+\nu-2,1} & 1 \\ \beta_{j,\nu} & \beta_{j+1,\nu-1} & \beta_{j+2,\nu-2} & \cdots & \beta_{j+\nu-2,2} & \beta_{j+\nu-1,1} \end{vmatrix}.$$

Denote this determinant by Δ_ν. Let $b_{n,m}$ be any set of positive numbers such that $b_{n,m} \geqq |\beta_{n,m}|$. We modify Δ_ν by first replacing $\beta_{n,m}$ by $b_{n,m}$ and then

[5] Proceedings of the Tokyo Physico-Mathematical Society, (3), vol. 13(1931), pp. 111–117.

agreeing to expand the resulting expression by the determinant rule except that every term is to be positive. If the resulting array is called D_ν, we have $|\Delta_\nu| \leqq D_\nu$. Expanding D_ν by the elements of its last column and putting $b_{n,m} = M_n\rho^{-m}$ we get

$$\begin{aligned}|c_{j+\nu}^{(j)}| &\leqq (M_{j+\nu-1} + 1)D_{\nu-1}\rho^{-1} \\ &\leqq (M_{j+\nu-1} + 1)(M_{j+\nu-2} + 1) \cdots (M_{j+1} + 1)M_j\rho^{-\nu} \\ &\leqq A(M + 1 + \epsilon)^\nu\rho^{-\nu}.\end{aligned}$$

From this estimate and the relations (19) and (17) it follows easily that

$$|c_i| \leqq A\rho^{-i}(M + 1 + \epsilon)^i + Br^{-i},$$

and the proof is complete.

6. Consider the linear space E of functions $f(x)$ defined as follows:

I. There exists in E a sequence of functions $\theta_0(x), \theta_1(x), \cdots$, all defined over a set S of real or complex numbers, which form a complete basis for E; that is, if $f(x)$ is in E, then

$$f(x) = \sum_{n=0}^{\infty} a_n\theta_n(x) \tag{20}$$

uniformly over some subset S' contained in S.

II. There is a positive number r such that for every $f(x)$ in E

$$a_n = O(r^{-n}). \tag{21}$$

III. There exists at least one set Γ in S such that if $g_m(x)$ and $f(x)$ are functions in E expressible over Γ by uniformly convergent series in the θ's, and if

$$f(x) = \sum_m g_m(x) \tag{22}$$

uniformly in Γ, then the expansion of $f(x)$ in terms of the θ's can be obtained by expanding each $g_m(x)$ and collecting the coefficients of like θ's.

Condition III is the analogue of the Weierstrass double series theorem. As an immediate consequence of Theorem V we have

THEOREM VI. *Let E be a space of functions satisfying conditions* I, II, *and* III, *and Γ a set satisfying* III. *Let $F_0(x), F_1(x), \cdots$ be in E and defined in Γ by the uniformly convergent series*

$$F_m(x) = \sum_{n=0}^{\infty} \beta_{m,n}\theta_{m+n}(x), \qquad \beta_{m,0} = 1, \tag{23}$$

where

$$|\beta_{m,n}| \leqq M_m\rho^{-n}, \qquad \limsup M_m = M < \infty. \tag{24}$$

Suppose that for x in Γ,

$$\limsup_{m\to\infty} | F_m(x) |^{1/m} < \xi \min\left(r, \frac{\rho}{1+M}\right), \tag{25}$$

for some positive number ξ less than 1. Then any function $f(x)$ which is in E and has a uniformly convergent expansion in terms of the θ's for x in Γ also has an expansion in terms of the F's:

$$f(x) = \sum_{m=0}^{\infty} c_m F_m(x), \tag{26}$$

which is uniformly convergent in Γ.

7. As a simple application of this theorem, we suppose that the θ's have the form

$$\begin{aligned} \theta_0(x) &= \Theta_0(x), \\ \theta_m(x) &= (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_m)\Theta_m(x) \quad (m = 1, 2, \cdots), \end{aligned} \tag{27}$$

where the functions $\Theta_m(x)$ are analytic and non-vanishing for $| x | \leqq \sigma$. The points $\alpha_1, \alpha_2, \cdots$ need not be distinct, but are confined to a region $| x | \leqq \zeta < \sigma$. It is assumed that the α's and Θ's are such that any function $f(x)$ which is analytic for $| x | \leqq \lambda$, $\lambda > \zeta$, can be expanded in a uniformly convergent series (20) for $| x | \leqq \rho_\lambda \leqq \lambda$, $\rho_\lambda > \zeta$, where ρ_λ is independent of $f(x)$. Thus, as far as condition I is concerned, we may choose for E the space E_λ of all functions analytic for $| x | \leqq \lambda$, S the set $| x | \leqq \sigma$, and S': $| x | \leqq \rho_\lambda$.

The θ's may be normalized so that

$$\Theta_m(\alpha_{m+1}) = 1. \tag{28}$$

We assume that this has been done. Let

$$\begin{aligned} \psi_m(x) &= \frac{\Psi_m(x)}{2\pi i(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_{m+1})}, \\ \Psi_m(x) &= 1 + k_{m,1}(x-\alpha_{m+1}) + \cdots + k_{m,m}(x-\alpha_2)\cdots(x-\alpha_{m+1}), \end{aligned} \tag{29}$$

with the k's to be determined. Then, if $m > n$, the product $\theta_m(x)\psi_n(x)$ will be analytic for $| x | \leqq \sigma$; and the integral

$$I_{m,n} = \int_C \theta_m(x)\psi_n(x)\,dx, \qquad C\colon\ | x | = \sigma,$$

will vanish. Obviously, we can orthogonalize $\psi_m(x)$ successively with respect to $\theta_{m-1}(x), \theta_{m-2}(x), \cdots, \theta_0(x)$ by choosing the k's successively in the order $k_{m,1}, k_{m,2}, \cdots, k_{m,m}$. This makes $I_{m,n} = \delta_{m,n}$ for all m and n; and the two sets $\{\psi_m(x)\}$ and $\{\theta_m(x)\}$ are biorthogonal.

We now make the additional assumption that the polynomials $\Psi_m(x)$ are uniformly bounded: $| \Psi_m(x) | \leqq K$ for $| x | \leqq \sigma$. Multiplying the uniformly

convergent series (20) by $\psi_m(x)$ and integrating term by term over $L: |x| = R$, $\zeta < R \leqq \rho_\lambda$, we get

(30) $$a_m = \int_L f(x)\psi_m(x)\,dx.$$

But, if $|x| > \zeta$, then

$$|\psi_m(x)| \leqq \frac{K}{2\pi(|x| - \zeta)^{m+1}}$$

and so

$$|a_m| \leqq NK\rho_\lambda(\rho_\lambda - \zeta)^{-m-1},$$

where N is the upper bound of $f(x)$ on $|x| = \rho_\lambda$. Thus E_λ satisfies condition II with $r = \rho_\lambda - \zeta$.

Let $f(x)$ and $g_m(x)$ be analytic for $|x| \leqq \lambda$ and let (22) hold uniformly for $|x| \leqq R \leqq \rho_\lambda$, $R > \zeta$. Then

$$g_m(x) = \sum_{n=0}^{\infty} b_{m,n}\theta_n(x)$$

uniformly in the region $|x| \leqq R$; and (20) holds uniformly in the same region. The coefficients a_n are given by (30), which, on combining with (22), reduces to

$$a_n = \sum_m \int_L g_m(x)\psi_n(x)\,dx.$$

Hence,

$$a_n = \sum_m b_{m,n},$$

and E_λ satisfies condition III if Γ is the set $|x| \leqq R$.

Choosing $F_m(x)$ as in (23) and (24), we have

$$F_m(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_m)H_m(x).$$

We assume that the functions $H_m(x)$ are analytic and uniformly bounded for $|x| \leqq \lambda$. The expansion (26) will then be valid for $|x| \leqq R$, $\zeta < R \leqq \rho_\lambda$, if the following inequality is satisfied for that same region:

$$\limsup\, [|x - \alpha_1|\,|x - \alpha_2| \cdots |x - \alpha_m|]^{1/m} < \xi \min\left(r, \frac{\rho}{1 + M}\right).$$

In particular, if $\zeta' = \limsup |\alpha_n|$, the expansion will be valid in the circle

$$|x| < \min\left(\rho_\lambda - \zeta, \frac{\rho}{1 + M}\right) - \zeta'$$

provided the right member of this equation is larger than ζ.

Starting with a set of θ's for which we already know the possibility of expanding $f(x)$ in form (20), we are led to conclude the possibility of expanding $f(x)$

in terms of the more general functions $F_m(x)$. Thus, starting with an initial set of θ's, we may be able to iterate the process so as to get successively more and more general results, by choosing at each stage the preceding F's as the new θ's.

Since much is known from other considerations about the possibility of expanding a function $f(x)$ in terms of Newton's interpolation series, we state in our final theorem the result of the above procedure for the special case where the initial expansions are of that form, i.e., the case $\Theta_m(x) = 1$.

THEOREM VII. *Let $\alpha_1, \alpha_2, \cdots$ be a bounded sequence of points in the complex plane. Let $\zeta \geqq |\alpha_n|$, $\zeta' = \limsup |\alpha_n|$. Suppose that any function $f(x)$ which is analytic for $|x| \leqq \lambda$, $\lambda_1 > \lambda > \lambda_0 \geqq 2\zeta + \zeta'$, has an expansion of the form*

$$f(x) = a_0 + \sum_{m=1}^{\infty} a_m(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m)$$

which is uniformly convergent for $|x| \leqq \rho_\lambda$, $\rho_\lambda = 2\zeta + \zeta' + \delta_\lambda$, $\delta_\lambda > 0$. Let

$$F_0(x) = 1 + h_0(x),$$

$$F_m(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m)[1 + h_m(x)] \quad (m = 1, 2, \cdots),$$

where $h_m(\alpha_{m+1}) = 0$, $h_m(x)$ is analytic and bounded in absolute value by N_m for $|x| \leqq \mu$, $\lambda_0 < \mu < \lambda_1$, $\limsup N_m = N < \infty$, and $\delta_\mu > M(\zeta + \zeta')$, $M = N\rho_\mu(\rho_\mu - \zeta)^{-1}$. Then $f(x)$ also has an expansion of the form

$$f(x) = \sum_{m=0}^{\infty} c_m F_m(x)$$

which is uniformly convergent for

$$|x| \leqq R < \zeta + \min\left[\delta_\lambda, \frac{\delta_\mu - M(\zeta + \zeta')}{1 + M}\right].$$

The proof parallels the general argument just given. The functions $h_m(x)$ can be expanded in the form

$$h_m(x) = \sum_{n=1}^{\infty} \beta_{m,n}\theta_{m,n}(x), \qquad \theta_{m,n}(x) = (x - \alpha_{m+1}) \cdots (x - \alpha_{m+n}).$$

For biorthogonal sets we have

$$\psi_{m,n}(x) = \frac{1}{2\pi i(x - \alpha_{m+1}) \cdots (x - \alpha_{m+n+1})}, \qquad \psi_n(x) = \psi_{0,n}(x).$$

Hence

$$\beta_{m,n} = \int_L h_m(x)\psi_{m,n}(x)\,dx,$$

$$|\beta_{m,n}| \leqq N_m\rho_\mu(\rho_\mu - \zeta)^{-n-1},$$

and the theorem follows.

For the particular case $\alpha_1 = \alpha_2 = \cdots = 0$, the hypotheses of the theorem are obviously satisfied for $\rho_\lambda = \lambda$, $0 < \lambda < \infty$; so we get Takenaka's theorem to the effect that any function $f(x)$ which is analytic for $| x | \leqq \lambda$ may be expanded in the uniformly convergent series

$$f(x) = \sum_{n=0}^{\infty} x^n[1 + h_n(x)], \qquad h_n(0) = 0,$$

in the region

$$| x | \leqq R < \min\left(\lambda, \frac{\mu}{1 + N}\right),$$

where $h_n(x)$ is analytic and bounded by N_n for $| x | \leqq \mu$ and $\limsup N_n = N$. Takenaka's theorem has been shown to include a large number of special expansions which occur in the literature,[6] of which the Neumann expansions in Bessel functions are best known.

INSTITUTE FOR ADVANCED STUDY.

[6] G. S. Ketchum, Transactions of the American Mathematical Society, vol. 40(1936), pp. 208–224.

TERNARY TRILINEAR FORMS IN THE FIELD OF COMPLEX NUMBERS

By R. M. Thrall and J. H. Chanler

1. **Introduction**. The specific problem of this paper is a classification of ternary trilinear forms with coefficients in the field of complex numbers. The complete solution is given by the table on page 689, and Theorem 12 gives necessary and sufficient conditions for the equivalence of two ternary trilinear forms under non-singular linear transformations on their sets of variables.

However, the authors consider the introduction of geometric methods to the study of trilinear forms of more importance than the specific results obtained. A trilinear form defines certain algebraic transformations between subspaces (manifolds) of the spaces of the variables. These subspaces may in general be curves, surfaces, or even isolated points; but for ternary trilinear forms they are plane cubic curves and the transformations between them are birational.

Consider the general trilinear form

$$F(x, y, z) = \sum a_{hij} x_h y_i z_j ,$$

where there are r_h x's, r_i y's, r_j z's and the a_{hij} are arbitrary complex numbers. Associated with the form is the three-way matrix (a_{hij}), called the matrix of the form. We suppose that the numbers r_h, r_i, r_j are the smallest numbers of x's, y's, z's, respectively, in terms of which the form can be expressed.[1] Two forms F and F' are called *equivalent*, and we write $F \sim F'$, if F can be sent into F' by non-singular linear transformations on the sets of variables taken separately. The totality of forms equivalent to a given form F is said to constitute a class of forms denoted by $[F]$. If $F \sim F'$, then $[F] = [F']$.

For a given three-way matrix (a_{hij}) there are six ways in which sets of variables x, y, z can be associated with the elements a_{hij} to produce trilinear forms.[2] Two trilinear forms derived from the same matrix are

$$F(x, y, z) = a_{111}x_1y_1z_1 + a_{112}x_1y_1z_2 + a_{113}x_1y_1z_3 + a_{121}x_1y_2z_1 + a_{122}x_1y_2z_2 + a_{123}x_1y_2z_3,$$

with $r_h = 1, r_i = 2, r_j = 3$, and

$$\begin{aligned} F(x, z, y) &= F'(x, y, z) \\ &= a_{111}x_1z_1y_1 + a_{112}x_1z_1y_2 + a_{113}x_1z_1y_3 + a_{121}x_1z_2y_1 + a_{122}x_1z_2y_2 + a_{123}x_1z_2y_3, \end{aligned}$$

Received January 17, 1938; presented to the American Mathematical Society, September 10, 1937.

[1] These numbers are equal to the h, i, j index ranks introduced by R. Oldenburger, IV. That these index ranks are the smallest numbers of variables in terms of which the form can be expressed was noted by H. R. Brahana, III, pp. 190–191. (Roman numerals refer to the bibliography at the end of paper.)

[2] VI, p. 384.

with $r_h = 1$, $r_i = 3$, $r_j = 2$. We say that two forms F and F' possess general equivalence and write $F \approx F'$ if $F \sim \bar{F}'$, where $\bar{F}'$ is one of the six forms associated with the matrix (a'_{hij}) of F'. The totality of forms generally equivalent (or g-equivalent) to a given form F are said to constitute a general class (or g-class) denoted by $\{F\}$.

The fundamental problem in the classification of trilinear forms is twofold: (1) the determination of necessary and sufficient conditions for equivalence (or g-equivalence) of two given trilinear forms; (2) the determination of a set $E(r_h, r_i, r_j)$ of forms which includes one and only one form from each class (r_h, r_i, r_j).[3] The solution of (2) obviously implies that of (1) although the converse is not true, as can be seen for the case of ternary trilinear forms by comparing Theorem 12 with the table of g-classes on page 689. (1) is listed as a separate problem, since it may have a simpler solution than (2), and is of some interest in itself.

If the r_ν are all different, the g-classes with invariants[4] r_h, r_i, r_j are determined by the classes for any particular ordering of the r_ν. In such cases we order the numbers so that $r_h > r_i > r_j$; then the set E which includes one and only one form F from each class (r_h, r_i, r_j) will also include one and only one form from each g-class with invariants r_ν. If, however, some of the numbers r_ν are equal, such a basis E for the classes of forms (r_h, r_i, r_j) may give more than one representative for certain g-classes, although it will still include at least one for each g-class. Evidently if a g-class has more than one representative in E, it has 2, 3, or 6. In general, to determine the g-classes with invariants r_ν, we first obtain the set E for some one of the orderings of the r_ν and then find what g-classes are thereby multiply represented and in what manner.

2. **Trilinear forms** (3, 3, 3); **general geometric theory**. It has been shown[5] that the problem of determining the classes of forms is abstractly identical with the problem of classifying the matrices[6]

$$M_x(x) = (x_{ij}) = \Big(\sum_h a_{hij} x_h\Big)$$

under multiplication on left and right by non-singular constant matrices and linear transformations on x. Instead of $M_x(x)$ we could use

$$M_y(y) = (y_{hj}) = \Big(\sum_i a_{hij} y_i\Big) \qquad \text{or} \qquad M_z(z) = (z_{hi}) = \Big(\sum_j a_{hij} z_j\Big).$$

[3] I.e., a class with two-way rank invariants r_h, r_i, r_j.

[4] The two-way rank invariants of any form considered as unordered numbers are invariants of the g-class to which the form belongs.

[5] VI, pp. 385–387.

[6] This matrix was introduced by H. R. Brahana (III, p. 196) and later named h-characteristic matrix by R. Oldenburger, V, p. 673. The use in V of variables ρ_h instead of x_h seems unnecessary in view of the meaning of this matrix as discussed in this paper and in VI, pp. 385–387.

The projective invariants of the ternary cubic $|M_z(x)| = X(x) = 0$ are therefore invariants of the trilinear form $F = \sum a_{hij}x_hy_iz_j$. Hence if $F \sim F'$, then $X(x) = 0$ is projectively equivalent to $X'(x) = 0$. Furthermore, any cubic projectively equivalent to $X(x) = 0$ will serve as $X'(x) = 0$ for at least one member F' of the class $[F]$. We may therefore choose one representative from each projective class (or case) of ternary cubics and insist that the member of $E(3, 3, 3)$ from any class $[F]$ have one of these canonical cubics (or a constant multiple of it) for $X(x)$. Any ternary cubic is projectively equivalent to one of the following:

(1) $$x_1^3 = 0;$$

(2) $$x_1^2x_2 = 0;$$

(3) $$x_1x_2(x_1 - x_2) = 0;$$

(4) $$x_1x_2x_3 = 0;$$

(5) $$x_3(x_1^2 + x_2x_3) = 0;$$

(6) $$x_1(x_1^2 + x_2x_3) = 0;$$

(7) $$x_1^3 - x_2^2x_3 = 0;$$

(8) $$x_1^3 + x_2^3 - x_1x_2x_3 = 0;$$

(9) an elliptic cubic;

(10) a cubic identically zero.

We shall say that $M_z(a)$ is an x-section of $M_z(x)$, the form F, and the matrix (a_{hij}) and define ξ, η, and ζ, respectively, to be the numbers (not necessarily finite) of x-, y-, and z-sections of rank one of a given form. We shall use the symbol $[\xi, \eta, \zeta]$ to describe the form.[7] We now prove some important preliminary theorems.

THEOREM 1. *If $M_z(a)$ is of rank one, then a is a multiple point of $X(x) = 0$.*

For the linear polar of $X(x)$ with respect to a can be written as the sum of determinants each involving two columns of $M_z(a)$. Since $M_z(a)$ is of rank one, each of these determinants vanishes.

THEOREM 2. *If for three collinear points $a^{(1)}, a^{(2)}, a^{(3)}$, $M_z(a^{(i)})$ is of rank one, then for all the points a on their line, $M_z(a)$ is of rank one.*

We normalize by taking $a^{(1)} = (1, 0, 0)$, $a^{(2)} = (0, 1, 0)$, and $a^{(3)} = (1, 1, 0)$. Then, since $M_z(1, 0, 0)$ and $M_z(0, 1, 0)$ are of rank one, we can find non-singular matrices P and Q such that

$$P \cdot M_z(x_1, x_2, 0) \cdot Q = \begin{pmatrix} x_1 & ax_2 & 0 \\ cx_2 & bx_2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

[7] The invariance of these and other numbers similarly defined was proved in VI, p. 386.

where $ac = 0$. Then for $M_z(1, 1, 0)$ to be of rank one we must have $b = 0$, and the theorem follows.

THEOREM 3. *If the determinant* $|M_z(x)|$ *vanishes identically,* $M_z(x)$ *cannot have more than two sections of rank one.*

The proof is similar to that for Theorem 2.

Immediate corollaries to these theorems are:

COROLLARY 1. *If* $X(x) = 0$ *is elliptic,* $M_z(x)$ *has no sections of rank one.*

COROLLARY 2. ξ *is finite unless* $X(x) = 0$ *is projectively equivalent to* $x_1^2x_2 = 0$, *or* $x_1^3 = 0$.

COROLLARY 3. *Unless* ξ *is infinite, it is* 0, 1, 2, *or* 3.

COROLLARY 4. *If* $\xi = 3$, $X(x) = 0$ *is projectively equivalent to* $x_1x_2x_3 = 0$, *and*

$$M_z(x) \sim \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}.$$

Given $F(x, y, z) = \sum a_{hij}x_hy_iz_j$, suppose that we ask for values $\bar{x}$ and $\bar{z}$ of x and z such that $F(\bar{x}, y, \bar{z}) \equiv 0$ in y. The condition for this is that the coefficients of y_1, y_2, y_3, which are bilinear forms in x and z, should vanish for $x = \bar{x}$, $z = \bar{z}$. Setting these coefficients equal to zero, and treating the result as three homogeneous linear equations in z_1, z_2, z_3, we see that the necessary and sufficient condition for a non-trivial solution in z is the vanishing of the determinant $M_z(x)$. Thus for each $\bar{x}$ such that $X(\bar{x}) = 0$, there will be at least one $\bar{z}$ for which $F(\bar{x}, y, \bar{z}) \equiv 0$. If $M_z(\bar{x})$ is of rank two, there will be just one such point $\bar{z}$ corresponding to $\bar{x}$; if $M_z(\bar{x})$ is of rank one, there will be two independent solutions $\bar{z}$, and therefore a line of points in the z-plane for which $F(\bar{x}, y, \bar{z}) \equiv 0$. $M_z(\bar{x})$ can not be of rank zero for $\bar{x} \neq (0, 0, 0)$; else r_h would be less than 3. The locus of $\bar{x}$ such that $M_z(\bar{x})$ is of rank two or less is given by $X(x) = 0$.

Since the conditions on $\bar{x}$, $\bar{z}$ for $F(\bar{x}, y, \bar{z}) \equiv 0$ are symmetric with respect to x and z, all the arguments concerning $M_z(x)$ apply to $M_x(z)$. In particular the locus of points $\bar{z}$ such that there exists a non-trivial $\bar{x}$ for which $F(\bar{x}, y, \bar{z}) \equiv 0$ is the curve $Z(z) = 0$.[8] Hence the trilinear form sets up a correspondence between the points of $X(x) = 0$ and those of $Z(z) = 0$. To consider this correspondence more precisely, let us return to the linear equations which determine $\bar{z}$. If $X(\bar{x}) = 0$, one of these three equations, say the third, is dependent on the other two. It is well known that two bilinear forms in x and z determine a quadratic Cremona transformation, say $z = z(x)$, between the x- and z-planes. As x runs through the points $\bar{x}$ of $X(x) = 0$, $\bar{z} = z(\bar{x})$ will run through a locus corresponding to $X(x) = 0$ under this quadratic transformation. But since $X(\bar{x}) = 0$, $F(\bar{x}, y, z(\bar{x})) \equiv 0$, hence $z(\bar{x})$ lies on $Z(z) = 0$.

[8] $Z(z) = |M_x(z)|$, $Y(y) = |M_y(y)|$.

Now if $\xi = \zeta = 0$, the correspondence between the points of $X(x) = 0$ and $Z(z) = 0$ will be (1, 1) without exception and will be given by the transformation $z = z(x)$. Instead of using the first two equations to define $z = z(x)$, we might use, say, the last two, obtaining another quadratic transformation, $z = z'(x)$, which coincides with $z = z(x)$ for $X(x) = 0$. We have therefore

THEOREM 4. *For $\xi = \zeta = 0$, $X(x) = 0$ and $Z(z) = 0$ are birationally related.*

The equations relating the points of $X = 0$ and $Z = 0$ can be found by setting $F(x, y^{(1)}, z) = F(x, y^{(2)}, z) = 0$, where $y^{(1)}$ and $y^{(2)}$ are any two different points y. Hence the trilinear form defines ∞^2 or a net of quadratic transformations which transform the points of $X(x) = 0$ in the same way, but which have no other common image. Since $\xi = 0$, there is no fundamental point common to all the members of this net. (Indeed the number of fundamental points common to all the members of the net is always precisely ξ.)

THEOREM 5. *For $\xi = \zeta = 0$, $X(x) = 0$ is projectively equivalent to $Z(z) = 0$.*

If $X(x) = 0$ is elliptic, so is the birationally related $Z(z) = 0$. But two elliptic cubics which are birationally related are projectively equivalent.[9] If $X(x) = 0$ is rational or degenerate, it follows from the analyticity of the transformations that $X = 0$, $Z = 0$ must have the same number of irreducible pieces; i.e., they must both be irreducible, or they must both factor into a line and a conic, or they must both factor into three lines. Furthermore, the transformations must be (1, 1) on the points of $X(x) = 0$, considered either as a point set or as points of an algebraic manifold; i.e., a nodal point counts once as a member of the point set, but twice as a point of the manifold. Hence, the number and nature of the singularities of $X = 0$ and $Z = 0$ must be the same. But if this is true, the curves are projectively equivalent. Finally if $X(x) \equiv 0$, the analyticity of the transformations insures that $Z(z) \equiv 0$.

An important consequence of this theorem is that by non-singular linear transformations on x and z we may insure that $X = 0$ and $Z = 0$ are constant multiples of one of the canonical cubics listed above. That this is possible follows from the fact that linear transformations on y and z replace M_x by PM_xQ and so replace $|M_x|$ by $|P|\cdot|Q|\cdot|M_x| = c\,|M_x|$. When X and Z are so transformed, the quadratic transformations become those of a cubic into itself. In the case $X \equiv 0$, $Z \equiv 0$, the algebraic transformation from x to z induced by the trilinear form is (1, 1) without exception throughout the projective plane. It must therefore be a projectivity of the plane into itself, and by a linear transformation on z we may insist that it be identity. Then since $F(x, y, x) \equiv 0$ in x and y, we have

$$M_y(y) \sim \bar{M}_y(y) = \begin{pmatrix} 0 & y_3 & y_2 \\ -y_3 & 0 & y_1 \\ -y_2 & -y_1 & 0 \end{pmatrix},$$

[9] I, p. 51.

which has no y-section of rank one; so $\eta = 0$. To show this, we write

$$F(x, y, z) \equiv \sum y_{hj} x_h z_j .$$

Then $M_y(y) = (y_{hj})$ is the matrix of a bilinear form in x and z. The condition $F(x, y, x) \equiv 0$ requires that $y_{hj} = -y_{jh}$; whence

$$M_y(y) = \begin{pmatrix} 0 & y_{12} & y_{13} \\ -y_{12} & 0 & y_{23} \\ -y_{13} & -y_{23} & 0 \end{pmatrix}.$$

Now saying that $F(x, y, z)$ can be expressed in terms of three and no fewer than three y's is equivalent to saying that all of the y_{hj} can be expressed in terms of some three of them. In this case y_{12}, y_{13}, y_{23} must be independent. Hence $M_y(y) \sim \bar{M}_y(y')$, where $y_3' = y_{12}$, $y_2' = y_{13}$, $y_1' = y_{23}$. Since $Y(y) \equiv 0$ and $M_y(y) \sim M_x(y) \sim M_z(y)$, we have

THEOREM 6. *There is just one g-class* $[0, 0, 0]$, $X(x) \equiv Y(y) \equiv Z(z) \equiv 0$, *and it is uniquely defined by the conditions* $\xi = \zeta = 0$, $X(x) \equiv 0$.

3. **Trilinear forms** (3, 3, 3) **with sections of rank one**. We now remove the restriction $\xi = \zeta = 0$.

THEOREM 7. *If* $M_x(a)$ *and* $M_z(a')$ *are sections of rank one, and if* $l(x) = 0$, $l'(z) = 0$ *are the lines in* x *and* z *corresponding to* $z = a'$, $x = a$, *respectively, then the line* $l(x) = 0$ *and the point* $x = a$ *are incident or not according as the line* $l'(z) = 0$ *and the point* $z = a'$ *are incident or not.*

The proof is similar to that for Theorem 2. We shall refer to Theorem 7 as the incidence theorem.

THEOREM 8. *If* $\xi = 3$, *then* $\eta = \zeta = 3$.

For we saw that

$$M_z(x) = \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}$$

represents the only class for $\xi = 3$. But then $M_y(x) = M_z(x) = M_x(x)$; so $\eta = \zeta = 3$.

THEOREM 9. *If* $\xi = \infty$, *then one of* η *and* ζ *is* ∞, *and the other is* 1 *or* 2, *there being just two g-classes in this case.*

We have seen that $\xi = \infty$ implies a line of double points in $X(x) = 0$. Let this line be $x_1 = 0$. Then by the argument of §2, we obtain

$$(1) \qquad M_z(x) \sim M_z^{(1)}(x) = (a_{ij})x_1 + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_2 + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_3 ,$$

or

$$(2)\qquad M_x(x) \sim M_x^{(2)}(x) = (a_{ij})x_1 + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_2 + \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_3.$$

(1) and (2) give different classes but the same set of g-classes. In (1) $\eta = \infty$, and in (2) $\zeta = \infty$. To establish the second part of the theorem, it is sufficient to show that (1) gives rise to two classes, with $\zeta = 1$ and $\zeta = 2$, respectively. We consider

$$M_z^{(1)}(z) = \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_1 & 0 & 0 \\ z_2 & 0 & 0 \end{pmatrix}.$$

Since $M_z^{(1)}(0, 0, 1)$ is of rank one, $\zeta \geqq 1$. The point at which $z_{12} = z_{13} = 0$ will give a section of rank one. If $\zeta = 1$, this point must be $(0, 0, 1)$; whence $a_{123} = a_{133} = 0$. Since z_3 must appear, $a_{113} \neq 0$. Then there exist non-singular matrices P and Q such that

$$PM_z^{(1)}Q = \begin{pmatrix} z_3 & z_1 & z_2 \\ z_1 & 0 & 0 \\ z_2 & 0 & 0 \end{pmatrix},$$

and we have only one class, whose rank invariants are $[\infty, \infty, 1]$. Similarly, if the point $z_{12} = z_{13} = 0$ is not $(0, 0, 1)$, we may take it to be $(1, 0, 0)$ giving $a_{121} = a_{131} = 0$, and

$$M_z^{(1)}(z) \sim \begin{pmatrix} 0 & z_3 & z_2 \\ z_1 & 0 & 0 \\ z_2 & 0 & 0 \end{pmatrix},$$

giving just one class $[\infty, \infty, 2]$. In both cases $Z(z) \equiv 0$.

Suppose $X \equiv 0$ and Y has no multiple line. Then unless $Y(y) \equiv 0$, it cannot possibly have enough pieces and double points to map into the whole x-plane. Hence there are exactly three g-classes for which any of X, Y, Z vanish identically: the two given by the theorem just preceding, and the third for which $\xi = \eta = \zeta = 0$, $X \equiv Y \equiv Z \equiv 0$.

Having disposed of the g-classes for any one of ξ, η, $\zeta > 2$, we now consider one of them equal to 2.

THEOREM 10. *The condition $\xi = 2$ implies that η, $\zeta \geqq 1$.*

We may evidently suppose that $M_x(1, 0, 0)$ and $M_x(0, 1, 0)$ are the x-sections of rank one. Then the line $x_3 = 0$, joining these double points, must be a factor of $X(x) = 0$ and corresponds to one or more sections of rank one in M_y and M_z.

4. **Determination of the g-classes** (3, 3, 3). Theorems 8, 9, 10 imply that for any form two of the numbers ξ, η, ζ are equal. We shall therefore get at least one member of each g-class if we list a representative of each class with $\xi = \zeta$. Hence we need consider only classes with the following rank invariants: [0, 0, 0], [0, 1, 0], [1, 0, 1], [1, 1, 1], [1, 2, 1], [2, 1, 2], [2, 2, 2], [3, 3, 3], $[\infty, 1, \infty]$, $[\infty, 2, \infty]$. We have already proved that there is just one class for each of the last three possibilities. We shall now consider the canonical cubics listed in §2; taking each in turn as $X(x)$, we shall determine for it the classes $[\xi, \eta, \xi]$.

For $X(x) = x_1^3$, and for $X(x) = x_1^2x_2$, the geometric methods used for the other cases break down due to the double and triple lines; we have therefore entered in the appended table of g-classes, for Cases 1 and 2, results obtained previously.[10]

Case 3. $X(x) = cx_1x_2(x_1 - x_2)$. Since $X(x) = 0$ has only one multiple point, $\xi = 0$ or 1. For $[0, \eta, 0]$ we may suppose $Z(x) = c_1X(x)$. To each piece of $X(x) = 0$ will correspond a piece of $Z(z) = 0$, and we can choose coördinates in z so that $x_1 = 0 \leftrightarrow z_1 = 0$; $x_2 = 0 \leftrightarrow z_2 = 0$; $x_1 - x_2 = 0 \leftrightarrow z_1 - z_2 = 0$. We recall that if two rational curves with parameters t and τ respectively are in birational correspondence, this correspondence can be expressed by a relation between the parameters: $\tau = (\alpha t + \beta)/(\gamma t + \delta)$.[11] Applying this theorem to $x_1 = 0$ and $z_1 = 0$ written respectively as $\rho x_1 = 0$, $\rho x_2 = 1$, $\rho x_3 = t_1$; $\rho z_1 = 0$, $\rho z_2 = 1$, $\rho z_3 = \tau_1$, we have $\tau_1 = (\alpha t_1 + \beta)/(\gamma t_1 + \delta)$. Since the triple points $x = (0, 0, 1)$ and $z = (0, 0, 1)$ must correspond, we have $t_1 = \infty \leftrightarrow \tau_1 = \infty$ or $\gamma = 0$. Similar arguments apply when we write $x_2 = 0$ as $\rho x_1 = 1$, $\rho x_2 = 0$, $\rho x_3 = t_2$, and $x_1 - x_2 = 0$ as $\rho x_1 = 1$, $\rho x_2 = 1$, $\rho x_3 = t_3$, and the same for z with τ_i instead of t_i. Thus we may write

$$\tau_i = \alpha_i t_i + \beta_i \qquad (i = 1, 2, 3). \tag{1}$$

Now for $x = x(t_i)$, $z = z(\tau_i) = z(\alpha_i t_i + \beta_i)$, $(i = 1, 2, 3)$, we have

$$F(t_i, y) \equiv F(x(t_i), y, z(\alpha_i t_i + \beta_i)) \equiv F(\bar{x}, y, z(\bar{x})) \equiv 0$$

in y and t_i. The automorphisms of $X(x)$ are generated by

$$A_1 : x_3' = ax_1 + bx_2 + x_3 ; \qquad A_2 : x_3' = cx_3 .$$

Under A_1, t_1 is replaced by $t_1' = t_1 + b$, t_2 by $t_2' = t_2 + a$, t_3 by $t_3' = t_3 + a + b$; under A_2, t_i is replaced by $t_i' = ct_i$ $(i = 1, 2, 3)$. We may evidently use these transformations on t_i and τ_i to modify (1) above. We get

$$\tau_1' = \alpha_1' t_1' ; \qquad \tau_2' = \alpha_2' t_2' ; \qquad \tau_3' = t_3' + \beta_3' , \quad (\beta_3' = 0 \text{ or } 1). \tag{1$'$}$$

Then the conditions

$$\begin{aligned} F(t_1', y) &\equiv y_{22} + y_{32}t_1' + y_{23}\tau_1' + y_{33}t_1'\tau_1' \\ &\equiv y_{22} + (y_{32} + \alpha_1' y_{23})t_1' + \alpha_1' y_{33} t_1'^2 \equiv 0, \\ F(t_2', y) &\equiv F(t_3', y) \equiv 0 \end{aligned} \tag{2}$$

[10] VI, p. 409.
[11] II, p. 140.

yield the respective sets of equations (when we write α_i, β_i for α_i', β_i'):

$$
\begin{aligned}
&y_{22} = y_{33} = y_{32} + \alpha_1 y_{23} = 0,\\
&y_{11} = y_{33} = \alpha_2 y_{13} + y_{31} = 0,\\
&y_{11} + y_{12} + y_{21} + y_{22} + \beta_3 y_{13} + \beta_3 y_{23}\\
&\qquad = y_{31} + y_{13} + y_{32} + y_{23} + \beta_3 y_{33} = y_{33} = 0.
\end{aligned} \tag{3}
$$

Hence

$$M_y = \begin{pmatrix} 0 & y_{12} & y_{13} \\ y_{21} & 0 & y_{23} \\ -\alpha_2 y_{13} & -\alpha_1 y_{23} & 0 \end{pmatrix},$$

where

$$y_{12} + y_{21} + \beta_3 y_{13} + \beta_3 y_{23} = 0, \qquad (1 - \alpha_2)y_{13} + (1 - \alpha_1)y_{23} = 0. \tag{4}$$

Unless equations (4) are dependent, there will not be the three independent y_{hj} which are necessary if $r_i = 3$. If $\alpha_2 = \alpha_1 = 1$, $\beta_3 = 0$, we have $| M_y | \equiv 0$. Hence $\beta_3 = 1$, and the sole class is represented by

$$M_y = \begin{pmatrix} 0 & y_1' & y_2' \\ -y_1' - y_2' - y_3' & 0 & y_3' \\ -y_2' & -y_3' & 0 \end{pmatrix},$$

where $y_1' = y_{12}$, $y_2' = y_{13}$, $y_3' = y_{23}$. We note that $\eta = 0$ as a consequence of $\xi = \zeta = 0$ and $X(x) = cx_1x_2(x_1 - x_2)$.

Next consider $[1, \eta, 1]$. Using the incidence theorem, we may suppose $x = (0, 0, 1) \to z_1 = 0$; $z = (0, 0, 1) \to x_1 = 0$; whence $y_{23} = y_{32} = y_{33} = 0$. Then $x_2 = 0$ and $x_1 - x_2 = 0$ must correspond to distinct lines in $Z(z) = 0$, and the intersection of these lines in z must correspond to the intersection of $x_2 = 0$ and $x_1 - x_2 = 0$; i.e., it must be on $z_1 = 0$. Since it is the only multiple point of $Z(z) = 0$, this intersection must be $z = (0, 0, 1)$, in accordance with our initial supposition that $M_z(0, 0, 1)$ was of rank one. Hence we may choose coördinates so that $Z(z) = c_1X(x)$, and so that $x_2 = 0 \leftrightarrow z_2 = 0$, $x_1 - x_2 = 0 \leftrightarrow z_1 - z_2 = 0$. Expressing this correspondence parametrically, we have $\tau_i = \alpha_i t_i + \beta_i$ $(i = 2, 3)$; by means of the automorphisms of X and Z, these transformations can be modified to $\tau_2 = t_2$, $\tau_3 = \alpha_3 t_3$. The conditions $F(t_2, y) \equiv F(t_3, y) \equiv 0$ give us $y_{11} = y_{13} + y_{31} = y_{11} + y_{12} + y_{21} + y_{22} = \alpha_3 y_{13} + y_{31} = 0$. These four relations imply that $r_i < 3$ unless $\alpha_3 = 1$. We have therefore the single class with

$$M_y = \begin{pmatrix} 0 & y_{12} & y_{13} \\ y_{12} - y_{22} & y_{22} & 0 \\ -y_{13} & 0 & 0 \end{pmatrix}, \qquad \eta = 2.$$

Case 4. $X(x) = cx_1x_2x_3$, where the lines are parametrically represented as $\rho x_1 = 0$, $\rho x_2 = 1$, $\rho x_3 = t_3$; $\rho x_1 = t_1$, $\rho x_2 = 0$, $\rho x_3 = 1$; $\rho x_1 = 1$, $\rho x_2 = t_2$, $\rho x_3 = 0$. For $[0, \eta, 0]$ we may suppose $x_i = 0 \leftrightarrow z_i = 0$ expressed by $\tau_i = \alpha_i t_i$ $(i = 1, 2, 3)$. The automorphism $x_i' = \rho_i x_i$ $(i = 1, 2, 3)$ produces the transformation

$$t_1' = \frac{\rho_1 t_1}{\rho_3}, \qquad t_2' = \frac{\rho_2 t_2}{\rho_1}, \qquad t_3' = \frac{\rho_3 t_3}{\rho_2},$$

so we may assume $\tau_i = \alpha t_i$, $\alpha = (\alpha_1\alpha_2\alpha_3)^{\frac{1}{3}}$. Imposing the condition $F(t_i, y) \equiv 0$, we have

$$M_y(y:\alpha) = \begin{pmatrix} 0 & y_{12} & -\alpha y_{31} \\ -\alpha y_{12} & 0 & y_{23} \\ y_{31} & -\alpha y_{23} & 0 \end{pmatrix}, \qquad \eta = 0;$$

$$M_y(y:\alpha) \sim M_y(y:\omega\alpha), \qquad (\omega^3 = 1).$$

Interchanging x and z, we get[12] $F(x, y, z: \alpha) \approx F(x, y, z: \alpha^{-1})$, since then $\tau' = t'\alpha^{-1}$. Interchanging y and z gives no new g-equivalence relations. Hence for $[0, \eta, 0]$, $X(x) = cx_1x_2x_3$, $F(x, y, z: \alpha) \approx F(x, y, z: \beta)$ if and only if $\beta^3 + \beta^{-3} = \alpha^3 + \alpha^{-3}$.

For $[1, \eta, 1]$ there are two cases possible: (1) with incidence; (2) without incidence. If (1) holds, arguments like those above show that there is just one g-class. It has $\eta = 1$, $Z(z)$ and $Y(y)$ in Case 4. Similarly if (2) holds, we have just one g-class, with $\eta = 2$, $Y(y)$ in Case 6 and $Z(z)$ in Case 4. For $[2, \eta, 2]$ we have just one g-class, with $\eta = 2$ and $Y(y)$ and $Z(z)$ in Case 4. We have already treated $[3, 3, 3]$.

The methods of investigation for Cases 5, 6, 7, 8 are those just applied to Cases 3 and 4. Hence we shall merely list the g-classes for these cubics in the appended table,[13] except for the two instances where $[\xi, \eta, \xi]$ and the case of $X(x)$ does not uniquely define the class. These instances are:

Case 6. $X(x) = cx_1(x_1^2 + x_2x_3)$, $[\xi, \eta, \xi] = [0, 0, 0]$,

$$M_y = \begin{pmatrix} (1-\alpha)y_{32} & y_{12} & y_{13} \\ -y_{12} & 0 & -\alpha y_{23} \\ -y_{13} & y_{32} & 0 \end{pmatrix},$$

where $\alpha \neq 1, 0$, and $F(x, y, z: \alpha) \approx F(x, y, z: \beta)$ if and only if $\beta = \alpha$ or α^{-1}; and

Case 8. $X(x) = c(x_1^3 + x_2^3 - x_1x_2x_3)$, with rank invariants $[0, 0, 0]$,

$$M_y = \begin{pmatrix} (\alpha^3 - 1)y_{23}/\alpha & y_{12} & -\alpha y_{31} \\ -\alpha y_{12} & (\alpha^3 - 1)y_{31}/\alpha & y_{23} \\ y_{31} & -\alpha y_{23} & 0 \end{pmatrix},$$

[12] I.e., the semicanonical form defined by $M_y(y:\alpha)$.

[13] Representatives of these g-classes are included in VI, pp. 397–411.

where $\alpha^3 \neq 1$, $F(x, y, z\colon \alpha) \approx F(x, y, z\colon \beta)$ if and only if $\alpha^3 + \alpha^{-3} = \beta^3 + \beta^{-3}$.

Before turning to the elliptic case we shall give an outline of the method followed in all the cases where $X(x)$ is rational and without a multiple factor. We have $\xi = \zeta$, where ξ cannot be greater than the number of linear factors of $X(x)$. The steps of our procedure are as follows: (1) we determine the projective case of $Z(z)$; (2) set up the correspondence between the parameters of the pieces of $X(x) = 0$ and $Z(z) = 0$; (3) normalize this correspondence by means of the automorphisms of $X(x)$ and $Z(z)$; (4) determine the y_{hj} so that $F(t, y) \equiv 0$, where t is the parameter of $X(x) = 0$, and so get M_y (here we rule out any correspondence set up in (2) which puts more than six independent conditions on the y_{hj}, since r_i must equal 3); (5) investigate the classes thus obtained for g-equivalence.

Case 9. $cX(x) = 4x_1^3 - g_2x_1x_3^2 - g_3x_3^3 - x_2^2x_3 = c_1Z(x)$. We give an analytic treatment based on the fact that the only birational transformations of a generic elliptic cubic into itself are: (1) the involutions interchanging the members of every pair of points collinear with some fixed point of the curve, and (2) the products of two such involutions.[14] We write $X(x) = 0$ parametrically as $\rho x_1 = \wp(u)$, $\rho x_2 = \wp'(u)$, $\rho x_3 = 1$, and do the same for $Z(z) = 0$, using parameter $\bar{u}$. The transformations from x to z are then given by $u \pm \bar{u} + a = 0$, where the plus sign goes with (1) and the minus sign with (2). This transformation is an automorphism (i.e., a collineation) if and only if a is a third-period. Since the automorphism $z_2' = -z_2$ replaces u by $-u$, we may always suppose our relation to be $u + \bar{u} + a = 0$. $F(x(u), y, z(\bar{u})) \equiv F(u, y)$ is an elliptic function of order six. Its vanishing identically in u and y is expressed by at most six conditions on the y_{hj}. That the six conditions are independent is seen by expanding $F(u, y)$ as a Laurent series about the poles $u = 0$ and $\bar{u} = -u - a = 0$ and equating the coefficients of negative powers of u and $\bar{u}$ and the constant term to zero. We get eight linear homogeneous equations in the y_{hj} whose matrix is of rank six. Denoting by $F(x, y, z\colon a)$ a trilinear form with $\bar{u} = -u - a$, we ask for what values of b is $F(x, y, z\colon a) \sim F(x, y, z\colon b)$? Such values of b must come from automorphisms, which can be expressed, as we have seen, by $u = \pm u' + \frac{1}{3}\Omega$, $\bar{u} = \pm \bar{u}' + \frac{1}{3}\Omega'$, where Ω, Ω' are any periods of $\wp(u)$, giving equivalence if and only if $b = \pm a + \frac{1}{3}\Omega$; i.e., for 18 values of b. Geometrically, this means that if the points of $X(x) = 0$ are grouped into sets of 18 in such a way that the sum or difference of the parameters of any two points in a set is a third-period, then there is one class for each such set. Interchanging the sets of variables gives no new g-equivalence. We have proved

THEOREM 11. *A class or g-class for which $X(x) = 0$ is a generic elliptic cubic*

[14] The word "generic" here excludes only the harmonic and equianharmonic cubics. To obtain the additional birational transformations of these cubics into themselves, we take certain collineations and the products of these collineations with each of the transformations (1) and (2) above (see II, pp. 151–155). Hence for these cases also we get the semicanonical form by considering only the involutions. However, new g-equivalences will be introduced among the points of the parallelogram of Theorem 11.

has two arithmetic invariants; one is the absolute invariant of $X(x) = 0$, *and the other is the group of values of a defining the set of points from which the center of the involution is obtained. If* ω_1, ω_2 *are the periods of* $\wp(u)$, *then there will be a class of forms* $F(x, y, z\colon a)$ *for each point a in the parallelogram whose vertices are* $0, \frac{1}{6}\omega_1, \frac{1}{6}\omega_1 + \frac{1}{3}\omega_2, \frac{1}{3}\omega_2$.

Case 10. $X(x) \equiv Y(y) \equiv Z(z) \equiv 0$. This case has already been considered.

TABLE OF G-CLASSES FOR $r_h = r_i = r_j = 3$

	[0,0,0]	[0,1,0]	[1,0,1]	[1,1,1]	[1,2,1]	[2,1,2]	[2,2,2]	[3,3,3]	[∞,1,∞]	[∞,2,∞]
(1)	(1)			(1)					(10)	
(2)	(2)			(2); (5)	(1)	(3)	(2)			(10)
(3)	(3)				(2)					
(4)	(4)*			(4)	(6)		(4)	(4)		
(5)	(5)	(7)	(3)							
(6)	(6)*	(8)	(4)							
(7)	(7)									
(8)	(8)*									
(9)	(9)*									
(10)	(10)									

The entry numbers on the left refer to the projective cases of $X(x)$, $Z(z)$ of the cubics in §2. The starred cases contain an infinite number of g-classes. There are no forms for the blocks without numbers. The numbers in the blocks give the projective case of $Y(y)$, there being just one class for each number except in the starred cases.

5. Conclusions and generalizations.

The results of the preceding section give the following

THEOREM 12. *A necessary and sufficient condition for the g-equivalence of two ternary trilinear forms F and F′ is*

(I) *that* ξ, η, ζ *equal* ξ', η', ζ' *in some order and* $X(x), Y(y), Z(z)$ *be projectively equivalent to* $X'(x), Y'(y), Z'(z)$ *in that same order; and*

(II) *that if* $\xi, \eta, \zeta = 0, 0, 0$ *and* $X(x)$ *is projectively equivalent to* (a) $x_1x_2x_3$, (b) $x_1(x_1^2 + x_2x_3)$, (c) $x_1^3 + x_2^3 - x_1x_2x_3$, *or* (d) $4x_1^3 - g_2x_1x_3^2 - g_3x_3^3 - x_2^2x_3$, *then in the semicanonical forms* $\bar{F}(x, y, z\colon \alpha)$ *and* $\bar{F}'(x, y, z\colon \beta)$ *of F and F′* (a) $\beta^3 + \beta^{-3} = \alpha^3 + \alpha^{-3}$, (b) $\beta + \beta^{-1} = \alpha + \alpha^{-1}$, (c) $\beta^3 + \beta^{-3} = \alpha^3 + \alpha^{-3}$ *or* (d) $\beta = \pm\alpha + \frac{1}{3}\Omega$, *where* Ω *is a period of* $\wp(u \mid g_2, g_3)$.

This theorem solves problem (1) for ternary trilinear forms. The above

table together with the semicanonical forms for (a), (b), (c), (d) gives a complete solution for problem (2).

As was indicated in the introduction, a major purpose of this paper was the application of well-developed geometric theory to trilinear form classification. Accordingly, we shall here indicate certain parts of the preceding theory which are susceptible to direct generalization. We shall limit ourselves in these generalizations to cases where $r_i = r_j$.

Then Theorem 1 becomes

THEOREM 1′. *If $M_z(a)$ is of rank $r_i - 2$, then a is a multiple point of the "surface" $X(x) = 0$.*

Theorem 2 remains true as stated. Corollaries 1, 2, 3 are replaced by

COROLLARY 1′. *ξ (the number of x-sections of rank $r_i - 2$) does not exceed the number of multiple points on $X(x) = 0$.*

If $r_h = r_i = r_j$, Theorem 4 becomes

THEOREM 4′. *For $\xi = \zeta = 0$, $X(x) = 0$ and $Z(z) = 0$ are birationally related.*

The direct analogues of Theorems 5 and 6 will not be true.

The theorems of §3 do not generalize directly, although there will be theorems of somewhat the same character to take their places.

BIBLIOGRAPHY

I. S. GANGULI, *The Theory of Plane Curves*, vol. 2, 1926.

II. SEVERI-LÖFFLER, *Vorlesungen über Algebraische Geometrie*, 1921.

III. H. R. BRAHANA, *Metabelian groups and trilinear forms*, this Journal, vol. 1(1935), pp. 190–191.

IV. R. OLDENBURGER, *On canonical binary trilinear forms*, Bulletin of the American Mathematical Society, vol. 38(1932), pp. 385–387.

V. R. OLDENBURGER, *Equivalence of multilinear forms singular on one index*, this Journal, vol. 2(1936), pp. 671–680.

VI. R. M. THRALL, *Metabelian groups and trilinear forms*, American Journal of Mathematics, vol. 40(1938), pp. 383–415.

UNIVERSITY OF MICHIGAN AND UNIVERSITY OF ILLINOIS.

EQUIDISTRIBUTION OF RESIDUES IN SEQUENCES

By Marshall Hall

1. **Introduction.** A sequence of rational integers $v_0, v_1, \cdots$ satisfying a rational integral linear recurrence

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_k u_n \tag{1.1}$$

is periodic[1] modulo an arbitrary prime p, that is,

$$u_{n+\tau} \equiv u_n \pmod{p} \tag{1.2}$$

for a fixed τ and all $n \geqq n_0(p)$. The polynomial $f(x) = x^k - a_1 x^{k-1} - \cdots - a_k$ is called the characteristic of (1.1). This paper is concerned with the distribution of the residues $0, 1, \cdots, p - 1 \pmod{p}$ in sequences satisfying (1.1). This distribution has been investigated in two papers, one by Ward[2] supposing $f(x)$ to be a cubic irreducible modulo p, and another by the author[3] supposing $f(x)$ to be any polynomial irreducible modulo p.

Here it is shown that if $f(x)$ is the product of a linear factor and an irreducible factor modulo p, the number of zeros in the different blocks will satisfy equations similar to those found in H. (Compare equations (2.3) in this paper with equations (13.8) in H.)

By a simple device these results may be used to show that when $f(x)$ is irreducible modulo p and the period of (v_n) is τ, an arbitrary residue $a \pmod{p}$ will occur $\tau p^{-1} + c_a$ times in any τ consecutive terms of (v_n), where $|c_a| < p^{\frac{1}{2}(k-1)}$.

The notation and terminology throughout are those of H.

2. **Distribution of zeros.** Suppose

$$f(x) \equiv (x - d)h(x) \pmod{p}, \tag{2.1}$$

where $h(x)$ is irreducible modulo p and at least of second degree.[4] A sequence

$$(v_n) \rightleftarrows g(x) \tag{2.2}$$

Received April 28, 1938.

[1] R. D. Carmichael, *On sequences of integers defined by recurrence relations*, Quarterly Journal of Mathematics, vol. 48(1920), pp. 343–372.

[2] Morgan Ward, *The distribution of residues in a sequence satisfying a linear recursion relation*, Transactions of the American Mathematical Society, vol. 33(1931), pp. 166–190.

[3] Marshall Hall, *An isomorphism between linear recurring sequences and algebraic rings*, Transactions of the American Mathematical Society, vol. 44 (1938), pp. 196–218. This paper will be referred to as H.

[4] For second order sequences see Marshall Hall, *Divisors of second order sequences*, Bulletin of the American Mathematical Society, vol. 43(1937), pp. 78–80, and the note to Theorem 13.5 of H.

will satisfy no recurrence modulo p of lower order than (1.1), if $g(x)$ is relatively prime[5] to the double modulus $p,f(x)$. Such a sequence is said to be *regular* modulo p. We shall investigate the distribution of zeros in the reduced cycles of the totality of regular sequences satisfying (1.1).

We note first that if (v_n) is any sequence satisfying a recurrence of characteristic $f(x)$, and if s is any integer prime to p, then $(s^n v_n)$ is a sequence satisfying a recurrence of characteristic $f_s(x) = s^k f(x/s)$. The zeros modulo p of $(s^n v_n)$ occur in the same positions as those in (v_n). Moreover, the reduced period of both sequences is the same. This is true for any $f(x)$, but will be used here only to simplify the proof of Theorem I.

THEOREM I. *If $f(x) \equiv (x-d)h(x) \pmod p$, where $h(x)$ is irreducible modulo p, then*

$$\sum_{i=1}^{\kappa} b_i = p^{k-2}, \tag{2.3.1}$$

$$\sum_{i=1}^{\kappa} b_i(b_i - 1) = (\mu - m)p^{k-3}, \tag{2.3.2}$$

$$\sum_{i=1}^{\kappa} b_i b_{i+r} = \begin{cases} \mu p^{k-3}, & r \not\equiv 0 \ (\kappa_1), \\ (\mu - m)p^{k-3}, & r \equiv 0 \ (\kappa_1), \end{cases} \tag{2.3.3}$$

$$(\mu, p-1) = m, \qquad p - 1 = mw, \qquad \mu = m\mu_1, \qquad \kappa = w\kappa_1, \tag{2.3.4}$$

where b_i is the number of zeros in a representative reduced cycle of the i-th block of regular sequences satisfying (1.1), *provided the κ blocks have been appropriately ordered.*

The subscripts in (2.3.3) are naturally taken modulo κ.

Under the secondary isomorphism

$$x^n g(x) \to v_n \tag{2.4}$$

the regular cycles (v_n) modulo p satisfying (1.1) form the cosets of the cyclic group $\{x\}$ in the group of residues prime to $[p,f(x)]$ (Theorem 13.1 of H).

Since $f(x) \equiv (x-d)h(x) \pmod p$, where $h(x)$ is irreducible modulo p, the group G of residues prime to $[p,f(x)]$ is representable as

$$G = r^i z^j \qquad (i = 0, \cdots, p-1; j = 0, \cdots, p^{k-1} - 2), \tag{2.5}$$

where r is a primitive root modulo p and hence of order $p-1$ and z is of order $p^{k-1} - 1$.

By definition of the reduced period μ

$$x^\mu \equiv b \ (\text{modd } p, f(x)), \tag{2.6}$$

where b is rational and

$$\mu\kappa = p^{k-1} - 1. \tag{2.7}$$

[5] See page 604 of Morgan Ward, *The arithmetical theory of linear recurring series*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 600–628.

Hence for $z_1 = z^c$ with c some integer prime to $p^{k-1} - 1$

$$z_1^\kappa \equiv sx \text{ (modd } p,f(x)), \tag{2.8}$$

where s is rational and

$$G = r^i z_1^j . \tag{2.9}$$

Hence the $(s^n v_n)$ cycles are obtained by taking every κ-th term of the (w_n) cycles, where $z_1^n \to w_n$ by the secondary isomorphism (Theorem 13.2 of H). Because of (2.9) there is but one z_1 block of reduced cycles prime to $[p,f(x)]$. Let (c_n) be a cycle of this block. The z_1 cycles are all regular except the cycle which is identically zero, the cycle whose $g(x) \equiv 0\ [p,h(x)]$ which contains no zeros, and the cycle $(x - d)z_1^n \to y_n$ which modd $p,f(x)$ is the same as the cycle z_1^n modd $p,h(x)$. We now apply Theorem 13.4 of H to the cycles (jc_n) $(j = 1, \cdots, p - 1)$ and (y_n). Here $n = 0, \cdots, p^{k-1} - 1$ and to obtain the same length we must take $p - 1$ reduced cycles of (y_n) since z_1^n (modd $p,h(x)$) has as its reduced period $(p^{k-1} - 1)(p - 1)^{-1}$ by Lemma 1 of Theorem 13.5 of H. By Theorem 13.4 of H for z_1^n modd $p,h(x)$ there are $(p - 1)[(p^{k-3} - 1)(p - 1)^{-1}] = p^{k-3} - 1$ pairs of zeros differing by a in position in (y_n) if $0 < a < p^{k-1} - 1$, $a \not\equiv 0\ (p^{k-1} - 1)(p - 1)^{-1}$ and $p^{k-2} - 1$ if $a \equiv 0\ (p^{k-1} - 1)(p - 1)^{-1}$. For all z_1 cycles modd $p,f(x)$ there are $p^{k-2} - 1$ pairs of zeros differing in position by a for $0 < a < p^{k-1} - 1$. Excluding the zeros in the (y_n) cycle and dividing by $p - 1$ we find that in (c_n) there are p^{k-3} pairs differing by a if $a \not\equiv 0\ (p^{k-1} - 1)(p - 1)^{-1}$ and none if $a \equiv 0\ (p^{k-1} - 1)(p - 1)^{-1}$. The terms $c_{n\kappa+i}$ $(n = 0, \cdots, \mu - 1)$ are a reduced cycle of the i-th block of $(s^n v_n)$ cycles, and these have, as remarked above, the same reduced period and position of zeros as (v_n) cycles. We now count the differences in positions of zeros in (c_n) in terms of the differences of positions of the cycles $(s^n v_n) = (c_{n\kappa+i})$. If there are b_i zeros in the i-th block, there will be $\sum_{i=1}^{\kappa} b_i(b_i - 1)$ differences in the same $(s^n v_n)$ cycle. But these are the differences $\equiv 0 \pmod{\kappa}$ in (c_n) and there are $\mu - m$ of these differences $\not\equiv 0\ ((p^{k-1} - 1)(p - 1)^{-1})$, where m is determined by (2.3.4). This yields the equation (2.3.2). Similarly, equations (2.3.3) enumerate the differences congruent to r modulo κ in terms of the b's. Equation (2.3.1) merely gives the total number of zeros in the (c_n) cycle.

THEOREM II. *Under the assumptions of Theorem* I, $b_i = p^{k-2}\kappa^{-1} + c_i$, *where*

$$|c_i| < \sqrt{\frac{\kappa - 1}{\kappa}} \sqrt{p^{k-2} - \frac{m\kappa + 1}{\kappa} p^{k-3}} < p^{\frac{1}{2}k-1}.$$

If we define quantities c_i by putting $b_i = p^{k-2}\kappa^{-1} + c_i$ in equations (2.3), we find

$$\begin{aligned} &\sum c_i = 0, \\ &\sum c_i^2 = p^{k-2} - \frac{(m\kappa + 1)}{\kappa} p^{k-3}, \end{aligned} \tag{2.10}$$

whence Theorem II follows immediately.

3. Distribution of arbitrary residues.

THEOREM III. *Let $f(x)$ be irreducible modulo p, and let the period of* (1.1) *modulo p be τ. Then the residue*[6] $a \not\equiv 0 \pmod p$ *will occur in a cycle of (u_n)* $n_a = \tau p^{-1} + c_a$ *times, where* $|c_a| < p^{\frac{1}{2}(k-1)}$.

If (u_n) satisfies (1.1), then $(v_n) = (u_n - a)$ satisfies a recurrence of characteristic $(x - 1)f(x)$ and n_a the number of residues congruent to a in a cycle (τ terms) of (u_n) will be the number of zero residues in τ (v_n) terms. But both the period and the reduced period of the (v_n) sequence are equal to τ. For $u_{n+\tau} \equiv u_n$ implies $v_{n+\tau} \equiv v_n$ and conversely, whence τ is the period of (v_n). The reduced period μ of (v_n) is the least solution of

$$x^{\mu} \equiv b \ (\text{modd } p, (x-1)f(x)), \tag{3.1}$$

whence $x^{\mu} \equiv b$ (modd p, $x - 1$), $1 \equiv b \pmod p$. Consequently,

$$x^{\mu} \equiv 1 \ (\text{modd } p, (x-1)f(x)) \tag{3.2}$$

and $\mu = \tau$ the period of (v_n).

Now Theorem II may be applied to the (v_n) cycle. n_a is the number of zeros in a (reduced) cycle of (v_n). Here (v_n) satisfies a recurrence of order $k + 1$ and

$$n_a = \frac{p^{k-1}}{\kappa} + c_a, \tag{3.3}$$

where $|c_a| < p^{\frac{1}{2}(k-1)}$. But here $\tau = \mu$ and $\mu\kappa = p^k - 1$, whence

$$n_a = \tau\left(\frac{p^{k-1}}{p^k - 1}\right) + c_a. \tag{3.4}$$

This gives

$$n_a = \frac{\tau}{p} + \frac{\tau}{p(p^k - 1)} + c_a.$$

Hence from Theorem II $n_a = \tau p^{-1} + c_a$, where $|c| < p^{\frac{1}{2}(k-1)}$.

Note that equations of the type

$$\sum c_i = 0, \qquad \sum c_i^2 = M \tag{3.5}$$

yield not only an upper bound on the magnitude of the c's $|c_i| \leqq \kappa^{-\frac{1}{2}}(\kappa - 1)^{\frac{1}{2}}M^{\frac{1}{2}}$ but also require that, for at least one c, $|c_i| \geqq \kappa^{-\frac{1}{2}}M^{\frac{1}{2}}$ (κ even), $|c_i| \geqq (\kappa - 1)^{-\frac{1}{2}}M^{\frac{1}{2}}$ (κ odd). If the c's are regarded as statistical variables, these equations give their mean and standard deviation.

[6] The number of zeros in (u_n) is given by Theorem 13.5 of H. According to this theorem in μ terms of (u_n) there are $\mu p^{-1} + c_0$ zeros where $|c_0| < p^{\frac{1}{2}k-1}$.

As an example of the application of Theorem III consider

$$u_{n+3} \equiv -u_{n+1} + u_n \pmod{5}. \tag{3.6}$$

$$\begin{array}{lll}
u_0 \equiv 0, & u_{11} \equiv 2, & u_{21} \equiv 4, \\
u_1 \equiv 0, & u_{12} \equiv 0, & u_{22} \equiv 3, \\
u_2 \equiv 1, & u_{13} \equiv 1, & u_{23} \equiv 4, \\
u_3 \equiv 0, & u_{14} \equiv 2, & u_{24} \equiv 1, \\
u_4 \equiv 4, & u_{15} \equiv 4, & u_{25} \equiv 4, \\
u_5 \equiv 1, & u_{16} \equiv 4, & u_{26} \equiv 3, \\
u_6 \equiv 1, & u_{17} \equiv 3, & u_{27} \equiv 2, \\
u_7 \equiv 3, & u_{18} \equiv 0, & u_{28} \equiv 1, \\
u_8 \equiv 0, & u_{19} \equiv 1, & u_{29} \equiv 1, \\
u_9 \equiv 3, & u_{20} \equiv 3, & u_{30} \equiv 1. \\
u_{10} \equiv 3, & &
\end{array}$$

Here the n_a will be equal to b_i of a recurrence of fourth order in which $p = 5$, $\mu = \tau = 31$, $k = 4$, $m = 1$, $\kappa = \kappa_1 = 4$, and the equations (2.3) for the distribution numbers are

$$\begin{aligned}
b_1 + b_2 + b_3 + b_4 &= 25, \\
b_1^2 + b_2^2 + b_3^2 + b_4^2 &= 175, \\
b_1b_2 + b_2b_3 + b_3b_4 + b_4b_1 &= 150, \\
b_1b_3 + b_2b_4 + b_3b_1 + b_4b_2 &= 150, \\
b_1b_4 + b_2b_1 + b_3b_2 + b_4b_3 &= 150.
\end{aligned} \tag{3.7}$$

Of these equations only three are independent and the solution associated with (3.6) is $b_1 = 9$, $b_2 = 3$, $b_3 = 6$, $b_4 = 7$. These are respectively the number of residues congruent to 1, 2, 4, and 3 in (3.6).

YALE UNIVERSITY.

ON THE FACTORIZATION OF GENERALIZED QUATERNIONS

By Gordon Pall

1. A fundamental theorem in the arithmetic of Lipschitz[1] integral quaternions

$$v = v_0 + i_1v_1 + i_2v_2 + i_3v_3, \tag{1}$$

where the v_i are rational integers and the i_α are the familiar Hamilton units ($i_\alpha^2 = -1$, etc.), is that any proper v (i.e., one in which $v_0, \cdots, v_3$ are coprime), whose norm $\sum v_i^2$ is divisible by an odd positive integer m, has exactly eight right-divisors t of norm m, these forming a class of left-associates

$$\pm t, \qquad \pm i_1t, \qquad \pm i_2t, \qquad \pm i_3t. \tag{2}$$

In this article a connection is set up between the problems of factoring "generalized quaternions" (defined in §3) and of representing the number 1 in a certain quaternary quadratic form S. Hence the problem is reduced to that of equivalence of quaternary quadratic forms. However, the order and genus of S is readily identified. Hence in all cases where there is but one class in this quaternary genus, a theorem of the type quoted above will follow; and when several classes occur in that genus, some similar theorem may be deducible.

Our definition of generalized quaternion, based on Hermite's identity,[2] connects the theory with ternary and quaternary quadratic forms, rather than with binary Hermitian forms as in Dickson's definition. For results similar to ours in Dickson's generalized quaternions, perhaps the best reference is *Ideals in generalized quaternion algebras*, Trans. Amer. Math. Soc., vol. 38(1935), pp. 436–446, by C. G. Latimer.

2. Our method is based on a process of Hermite's,[3] who in turn was guided by Gauss's algorithm for reducing the representation of numbers in a binary quadratic form to the solution of a quadratic congruence and to identifying the class of a form constructed from the solution. We shall introduce the method by exhibiting a similar device for quadratic fields. We shall confine ourselves, however, to fields in which the integers are of the form

$$v = v_0 + v_1\omega, \qquad v_0 \text{ and } v_1 \text{ rational integers}, \tag{3}$$

where $\omega^2 = -D$ is a non-square rational integer. There is no difficulty in extending the theory to $\omega^2 + \omega + \frac{1}{4}(1 - \Delta) = 0$. Similarly, in this article we

Received May 18, 1938.

[1] R. Lipschitz, Jour. de Math., (4), vol. 2(1886), French translation by J. Molk.

[2] C. Hermite, Jour. für Math., vol. 47(1854), pp. 313–330; *Oeuvres*, vol. 1, 1905, pp. 200–220, especially p. 212.

[3] Hermite, Jour. für Math., vol. 47(1854), pp. 343–345; *Oeuvres*, vol. 1, pp. 234–237.

shall treat only the simpler case of quaternions associated with Hermite's identity.[4] The extension by means of Brandt's generalization[5] of Hermite's identity is being investigated by students of the writer.

Let p be an odd prime not dividing D. How many divisors of norm p does v possess? For this we must consider

$$v_0 + v_1\omega = (u_0 + u_1\omega)(t_0 + t_1\omega), \qquad t_0^2 + Dt_1^2 = p. \tag{4}$$

On taking norms we see that p must divide $Nv = v_0^2 + Dv_1^2$. Assuming this and multiplying both sides of (4) by $\bar{t} = t_0 - t_1\omega$, we get

$$(v_0 + v_1\omega)(t_0 - t_1\omega) \equiv 0 \qquad \pmod{p},$$

$$v_0t_0 + Dv_1t_1 \equiv 0, \qquad v_1t_0 - v_0t_1 \equiv 0, \qquad \pmod{p}. \tag{5}$$

Now assume that p does not divide both (and hence either) of v_0, v_1. Then the condition $p \mid Nv$ makes either of conditions (5) imply the other. Either of them reduces to $t_0 \equiv et_1 \pmod{p}$, where e is an integer $\equiv v_0/v_1 \pmod{p}$.

Thus if $p \mid Nv$, (5) will be satisfied if and only if

$$\begin{aligned} t_0 &= pX_0 + eX_1, \\ t_1 &= X_1, \end{aligned} \tag{6}$$

in integers X_0, X_1; and then $v\bar{t} \equiv 0 \pmod{p}$, $v\bar{t} = pu$, $v\bar{t}t = put$, $v = ut$, provided for the last step $Nt = \bar{t}t = t_0^2 + Dt_1^2 = p$. If we substitute from (6), the condition $Nt = p$ becomes

$$pX_0^2 + 2eX_0X_1 + fX_1^2 = 1, \tag{7}$$

where $f = (e^2 + D)/p$. The form $[p, 2e, f]$ is of determinant D, and will represent 1 if and only if the prime p happens to be represented in $x_0^2 + Dx_1^2$; and then the number of divisors t of norm p of v will be 2 if $D > 1$, and 4 if $D = 1$; if $D > 1$, the divisors are $\pm t$, if $D = 1$ they are $\pm t$ and $\pm it$. When there is only one class in each genus, the condition for p to be presented in $x_0^2 + Dx_1^2$ can be expressed simply in terms of Legendre symbols.

3. With a symmetric matrix $a = (a_{\alpha\beta})$ of order 3 in a field $\mathfrak{F}$, we associate a system of quaternions

$$t = t_0 + i_1t_1 + i_2t_2 + i_3t_3, \tag{8}$$

where the t_i range over $\mathfrak{F}$ and i_1, i_2, i_3 satisfy the following multiplication table, $A_{\alpha\beta}$ denoting the cofactor of $a_{\alpha\beta}$ in a:

$$\begin{aligned} i_\alpha^2 &= -A_{\alpha\alpha} \qquad\qquad (\alpha = 1, 2, 3). \\ i_2i_3 &= -A_{23} + a_{11}i_1 + a_{12}i_2 + a_{13}i_3, \\ i_3i_2 &= -A_{32} - a_{11}i_1 - a_{12}i_2 - a_{13}i_3, \end{aligned} \tag{9}$$

[4] See footnote 2.

[5] H. Brandt, Math. Annalen, vol. 91(1924), pp. 308–309.

$i_3 i_1$, etc., being obtained by permuting subscripts cyclically. Multiplication by scalars (in $\mathfrak{F}$) is defined in the obvious way, addition by adding corresponding coördinates. Addition is commutative and associative, and scalar factors can be taken in and out of products. Multiplication is distributive over addition. In general, $ut = v$ is given by the formulas (which are essentially those of Hermite):[6]

$$
\begin{aligned}
v_0 &= u_0 t_0 - \sum_{\alpha,\beta=1}^{3} A_{\alpha\beta} u_\alpha t_\beta, \\
v_\alpha &= u_0 t_\alpha + u_\alpha t_0 + \sum_{\beta=1}^{3} w_\beta a_{\beta\alpha} \qquad (\alpha = 1, 2, 3),
\end{aligned}
\tag{10}
$$

$$w_1 = u_2 t_3 - u_3 t_2, \qquad w_2 = u_3 t_1 - u_1 t_3, \qquad w_3 = u_1 t_2 - u_2 t_1.$$

Seeing that multiplication is associative reduces to verifying

$$(i_\alpha i_\beta) i_\gamma = i_\alpha (i_\beta i_\gamma) \tag{11}$$

for all choices of subscripts 1, 2, 3. Forming the products by means of (9), we readily find

$$
\begin{aligned}
(i_\alpha i_\alpha) i_\beta &= i_\alpha (i_\alpha i_\beta) = -A_{\alpha\alpha} i_\beta, \\
i_\alpha (i_\beta i_\beta) &= (i_\alpha i_\beta) i_\beta = -A_{\beta\beta} i_\alpha, \\
(i_\alpha i_\beta) i_\alpha &= i_\alpha (i_\beta i_\alpha) = -2A_{\alpha\beta} i_\alpha + A_{\alpha\alpha} i_\beta, \\
(i_1 i_2) i_3 &= i_1 (i_2 i_3) = -\Delta - A_{23} i_1 + A_{31} i_2 - A_{12} i_3, \\
(i_2 i_1) i_3 &= i_2 (i_1 i_3) = \Delta + A_{32} i_1 - A_{13} i_2 - A_{21} i_3,
\end{aligned}
\tag{12}
$$

$i_2 i_3 i_1$, etc., being obtained by cyclic permutation. Here $\Delta = | a_{\alpha\beta} |$.

We define conjugates by

$$\bar{t} = t_0 - i_1 t_1 - i_2 t_2 - i_3 t_3, \tag{13}$$

so that by (9), $i_3 i_2 = \overline{i_2 i_3}$, and readily obtain

$$\bar{t} t = t \bar{t} = t_0^2 + \sum A_{\alpha\beta} t_\alpha t_\beta, \tag{14}$$

which we call the *norm of* t, or Nt. Since $\overline{i_\alpha i_\beta} = (-i_\beta)(-i_\alpha) = \bar{i}_\beta \bar{i}_\alpha$, and $\overline{t + u} = \bar{t} + \bar{u}$, we see that the conjugate of $(u_0 + \sum u_\alpha i_\alpha)(t_0 + \sum t_\beta i_\beta)$ is $(t_0 - \sum t_\beta i_\beta)(u_0 - \sum u_\alpha i_\alpha)$, that is,

$$\overline{ut} = \bar{t} \bar{u}. \tag{15}$$

Hence from $v = ut$ follows $\bar{v} = \bar{t}\bar{u}$, whence $v\bar{v} = ut\bar{t}\bar{u}$, or

$$Nv = Nu \cdot Nt. \tag{16}$$

This is Hermite's identity.

[6] See footnote 2.

4. In case the $a_{\alpha\beta}$ are rational integers we define an *integral quaternion* by (8) with rational integral t_i. By (9) the sum and product of integral quaternions are integral. If $v = ut$ in integral quaternions, we call t a *right-divisor* of v. Then by (16), $Nt \mid Nv$.

If t is a right-divisor of v, $\bar{t}$ is a left-divisor of $\bar{v}$. If $t_1 = t_2 = t_3 = 0$ so that t is a rational integer, t is a right-divisor of v if and only if $t \mid v_i$ $(i = 0, 1, 2, 3)$, and we can write $t \mid v$.

For a given v such that $p \nmid v$ but $p \mid Nv$, how many right-divisors of norm p does v possess? From

$$v = ut, \qquad Nt = p \tag{17}$$

follows $v\bar{t} = up$, $v\bar{t} \equiv 0 \pmod p$, or (cf. (10)) the system of homogeneous, linear congruences in t_0, t_1, t_2, t_3 with the matrix W obtained from

$$V \equiv \begin{bmatrix} 0 & \sum A_{1\alpha}v_\alpha & \sum A_{2\alpha}v_\alpha & \sum A_{3\alpha}v_\alpha \\ v_1 & a_{31}v_2 - a_{21}v_3 & a_{11}v_3 - a_{31}v_1 & a_{21}v_1 - a_{11}v_2 \\ v_2 & a_{32}v_2 - a_{22}v_3 & a_{12}v_3 - a_{32}v_1 & a_{22}v_1 - a_{12}v_2 \\ v_3 & a_{33}v_2 - a_{23}v_3 & a_{13}v_3 - a_{33}v_1 & a_{23}v_1 - a_{13}v_2 \end{bmatrix}$$

by adding $v_0, -v_0, -v_0, -v_0$ to the elements of the principal diagonal. In general if $p \mid Nv$, W is of rank 3, (mod p). If $p \mid v_0$, W becomes V and the crux of our method lies in

LEMMA 1. *If $p \mid v_0$, $p \mid Nv = \sum A_{\alpha\beta}v_\alpha v_\beta$, and $p \nmid \Delta v$, then two and at most two rows in V are linearly independent* (mod p).

To show that two rows are independent, it suffices since $p \nmid v$ to show that one of $\sum A_{\beta\alpha}v_\alpha$ is prime to p; if this were not so, p would divide

$$\sum_\beta a_{\beta\gamma} \sum_\alpha A_{\beta\alpha}v_\alpha = \Delta v_\gamma \qquad (\gamma = 1, 2, 3),$$

contrary to hypothesis. For the rest, it suffices to show that every three columns are linearly dependent. For the second, third, and fourth columns we use the multipliers v_1, v_2, and v_3, and adding obtain $(Nv, 0, 0, 0)$. For the second, third, and first columns we multiply by $\sum A_{2\alpha}v_\alpha$, $-\sum A_{1\alpha}v_\alpha$, and Δv_3, and on adding obtain $(0, a_{31}Nv, a_{32}Nv, a_{33}Nv)$. If these multipliers are all 0 (mod p), a glance at V shows that the second and third columns are proportional.

COROLLARY 1. *With the hypotheses of Lemma* 1, *if of the conditions*

$$\begin{aligned} &(\textstyle\sum A_{1\alpha}v_\alpha)t_1 + \cdots \equiv 0, \qquad v_1t_0 + (a_{31}v_2 - a_{21}v_3)t_1 + \cdots \equiv 0, \\ &v_2t_0 + (a_{32}v_2 - a_{22}v_3)t_1 + \cdots \equiv 0, \qquad v_3t_0 + \cdots \equiv 0, \quad (\mathrm{mod}\ p), \end{aligned} \tag{18}$$

the first is satisfied together with the α-th one of the others, where $p \nmid v_\alpha$, then all four conditions are satisfied. Further, for all such t_i,

$$p \mid Nt = t_0^2 + \sum A_{\alpha\beta}t_\alpha t_\beta. \tag{19}$$

To secure (19) we need only observe that $p \mid v\bar{t}$, $p \mid v\bar{t}t$, $p \mid v(Nt)$, $p \mid Nt$.

Conversely, any solution t of (18) satisfying $Nt = p$ is a right-divisor of v. For, $v\bar{t} = pu$, $v\bar{t}t = put$, $v = ut$.

For definiteness we can suppose that $\sum A_{1\alpha} v_\alpha$ is prime to p, and solve (18) in the form

$$t_0 \equiv at_2 + bt_3, \qquad t_1 \equiv ct_2 + dt_3, \qquad (\mathrm{mod}\ p), \tag{20}$$

where a, b, c, d are certain integers. Then (18) is equivalent to

$$\begin{aligned} t_0 &= pX_0 + aX_2 + bX_3, \qquad & t_2 &= X_2, \\ t_1 &= pX_1 + cX_2 + dX_3, \qquad & t_3 &= X_3, \end{aligned} \tag{21}$$

in integers X_i. Substituting these expressions in $t_0^2 + \sum A_{\alpha\beta} t_\alpha t_\beta = p$, we get

$$\sum_{i,j=0}^{3} r_{ij} X_i X_j = p. \tag{22}$$

By (19), $p \mid \sum r_{ij} X_i X_j$ for all integers X_i, whence p divides every r_{ij}; set $r_{ij} = ps_{ij}$. Thus (22) reduces to the equation

$$\sum s_{ij} X_i X_j = 1. \tag{23}$$

The number of divisors equals the number of solutions of (23).

Let P denote the matrix of transformation (21), which has determinant p^2; let q, s denote the matrices of the forms $Q = t_0^2 + \sum A_{\alpha\beta} t_\alpha t_\beta$, $S = \sum s_{ij} X_i X_j$. That the determinants of q and s are equal follows from

$$ps = P'qP. \tag{24}$$

We shall see that the orders of Q and S coincide. Since the index is unaltered by the real transformation (21), this will follow if we show for $k = 1, 2$, and 3 that the g.c.d. of the principal minor determinants and the doubles of the non-principal minor determinants of order k is the same for q and s.[7] Let σ denote a subsequence of k elements of (1234); for any matrix R let $R[\sigma_1\sigma_2]$ denote the minor determinant whose rows have the positions indicated by σ_1, and the columns those of σ_2. Then by (24) and a simple determinantal theorem,

$$p^k s[\sigma_1\sigma_2] = \sum_{\sigma,\sigma'} q[\sigma\sigma'] P[\sigma\sigma_1] P[\sigma'\sigma_2], \tag{25}$$

where the summation ranges over the $({}^4C_k)^2$ pairs σ, σ'. Since $p \nmid \Delta$, p cannot divide every $q[\sigma\sigma']$; if $\sigma_1 = \sigma_2$, the terms with $q[\sigma\sigma']$ and $q[\sigma'\sigma]$ are equal; hence the g.c.d. of the $q[\sigma\sigma]$ and $2q[\sigma\sigma']$ divides every $s[\sigma\sigma]$ and $2s[\sigma\sigma']$. The converse follows on solving (24) for q in terms of s.

In terms of the order invariants[8] recently introduced by the writer, let the divisor and o-invariants of the ternary form $f = \sum a_{\alpha\beta} x_\alpha x_\beta$ be d_1, o_1, and o_2. Since the $a_{\alpha\beta}$ are assumed integral, d_1 is even if o_1 is odd. The order invariants

[7] G. Pall, Quart. Jour. of Math., vol. 6(1935), pp. 30–51, Theorem 2.

[8] Same as footnote 7, Theorem 1.

of the primitive form Q, and therefore of S, are found to be $(o_1 d_1^2, o_2, o_1)$. Denote the primitive concomitants by $Q_1 = Q, Q_2, Q_3$. The genus of Q is characterized by the values of the principal characters $(Q_k \mid \omega_k)$ for the odd primes ω_k dividing $o_1 d_1^2, o_2, o_1$, respectively, and possibly certain supplementary characters $(-1|Q_k)$, $(2|Q_k)$, or $(-2|Q_k)$.[9] It is evident that if $(Q_3|\omega)$ is a character the same is true of $(Q_1|\omega)$. It is not difficult to verify also, by using a canonical form[10]

$$2^{\beta_1}m_1x_1^2 + 2^{\beta_2}m_2x_2^2 + 2^{\beta_3}m_3x_3^2 \quad \text{or} \quad 2^{\beta_1}m_1x_1^2 + 2^{\gamma}(mx_2^2 + m'x_2x_3 + nx_3^2) \pmod{2^h}$$

for f, that if any of $(-1|Q_3)$, $(2|Q_3)$, or $(-2|Q_3)$ is a character, the same is true of Q_1. It should be observed that the simultaneous characters need not be considered here, since for forms in less than five variables their values are fixed by those of the others.[11] Further, since Q represents 1, the value of any character due to Q_1 is $+1$.

Finally, it is evident from the process by which S was derived (or from (25)) that if S_k represents n, then Q_k represents $p^k n$. Hence the characters due to S_2 and Q_2 are equal, while the values of those of S_1 (or S_3) are obtained from those of Q_1 (or Q_3) by multiplying by the quadratic character of p. The form S will therefore not represent 1 unless for each character $(Q_1|\omega)$, $(-1|Q_1)$, $(2|Q_1)$, and $(-2|Q_1)$ which may happen to be an invariant of Q, the value obtained on substituting p for Q_1 is $+1$. When this holds, we may say that *p is consistent with the genus of Q.*

Integral quaternions of norm 1 are called *units*. Their number, possibly infinite, is equal to the number ρ of solutions of

$$t_0^2 + \sum A_{\alpha\beta}t_\alpha t_\beta = 1. \tag{25'}$$

If θ denotes an arbitrary unit, the ρ quaternions θt are called *left-associates*. If t is a right-divisor of v, the left-associates θt form ρ right-divisors of the same norm. We have now proved the following theorem in the case $p \mid v_0$, and shall remove this restriction in §5:

THEOREM 1. *Let $(a_{\alpha\beta})$ be a symmetric matrix of order 3, the $a_{\alpha\beta}$ integers, $(A_{\alpha\beta})$ the adjoint. Set $Q = t_0^2 + \sum A_{\alpha\beta}t_\alpha t_\beta$. Let p be an odd prime not dividing $|a_{\alpha\beta}|$, v an integral quaternion of the type defined in §3 and the beginning of §4, $p \mid Nv$, $p \nmid v$. Then v has no right-divisors of norm p unless p is consistent with the genus of Q, and then the number of right-divisors is equal to the number of representations of 1 in a certain form of the same genus as Q. If this genus contains but one class, there are exactly ρ right-divisors of norm p, these forming a class of left-associates.*

COROLLARY. *The last sentence of the theorem holds with p replaced by m, where m is a product of primes each consistent with the genus of Q, and*

(26) *m is prime to 2Δ, $m \mid Nv$, and m, v_0, v_1, v_2, v_3 are coprime.*

[9] H. J. S. Smith, *Coll. Math. Papers*, I, pp. 513–514.

[10] Same as footnote 7, Lemma 3.

[11] H. J. S. Smith, loc. cit., p. 515 (relation (A)).

We prove this by induction assuming the theorem as stated. Assume the corollary to be true for products m of h or fewer primes consistent with the genus of Q and dividing neither 2Δ nor v. Let p be such a prime.

Existence. From $v = ut$, $Nt = m$, $pm \mid Nv = NuNt$, follow $p \mid Nu$, $u = wx$, where $Nx = p$, $v = w(xt)$, $N(xt) = pm$.

Uniqueness. If $v = ux = wy$, $Nx = Ny = pm$, then $x = at$ and $y = bt'$, where $Nt = Nt' = m$, t and t' are left-associates since they are both right-divisors of v; by absorbing the unit factor on the left we can make $t = t'$; thus $ua = wb$, where a and b are both of norm p and hence are left-associates. Consequently, $x = at$ and $y = bt$ are left-associates.

There are no characters $(Q_1|\omega)$, $(-1|Q_1)$, $(2|Q_1)$, $(-2|Q_1)$ if and only if

$$\begin{aligned} &d_1 = 1,\ o_1 = 4 \text{ or } 8,\ 16 \nmid o_2; \text{ or} \\ (27)\qquad &d_1 = 1,\ o_1 = 4,\ 16 \mid o_2,\ (-1|f_2) = -1; \text{ or} \\ &d_1 = 2,\ o_1 = 1, \end{aligned}$$

f_2 denoting the reciprocal of $f = \sum a_{\alpha\beta}x_\alpha x_\beta$. *In these cases then, if there is but one class in the genus of Q, there will be exactly ρ right-divisors of norm m, for any positive m satisfying* (26). An attempt to extend the theorem to such an m in any case seems to fail because of the difficulty of securing that some minor determinant of order 2 in V is prime to m.

We may note that if m were negative we would have -1 at the right of the equation corresponding to (23), since we would divide by $|m|$ to keep the index unaltered. It may be worth while observing also that the chain of leading principal minor determinants in the matrix of (23) is p, $A_{11}p^2$, $pa_{33}\Delta$, Δ^2; which are independent of v.

5. The problem is reduced to the case $p \mid v_0$ by two lemmas:

LEMMA 2. *If v and w are integral quaternions, and m an integer prime to Nw, then v and wv have the same right-divisors of norm m.*

For if $v = ut$, $wv = (wu)t$. Conversely, if $wv = ut$, and $Nt = m$ is prime to $Nw = k$, then $kv = (\bar{w}u)t$, $kv\bar{t} = \bar{w}um$,

$$v\bar{t} = \bar{w}u(m/k) = \text{integral quaternion}.$$

Hence the coördinates of $\bar{w}u$ are each divisible by k, $v = (\bar{w}u/k)t$.

LEMMA 3. *Let p be an odd prime not dividing Δ, nor all of v_1, v_2, v_3. Then we can find a pure quaternion w such that Nw is prime to p but the real part of wv is divisible by p.*

It is clear that the condition on v_1, v_2, v_3 is satisfied if $p \mid Nv$, $p \nmid v$. I am indebted to R. E. O'Connor for the following simple proof of this lemma. The problem is to find integers w_1, w_2, w_3 to satisfy

$$(28)\qquad \begin{aligned} &(\textstyle\sum A_{1\beta}v_\beta)w_1 + (\sum A_{2\beta}v_\beta)w_2 + (\sum A_{3\beta}v_\beta)w_3 \equiv 0, \\ &\textstyle\sum A_{\alpha\beta}w_\alpha w_\beta \not\equiv 0, \qquad (\text{mod } p). \end{aligned}$$

Using matrix notation, we can write these in the form

$$(28') \qquad w'Av \equiv 0, \qquad w'Aw \not\equiv 0, \qquad (\text{mod } p),$$

where A is a symmetric matrix of order 3 with integer elements, of determinant prime to p, v and w are vertical vectors and the prime denotes transposition. By a sequence of elementary transformations we can find a matrix C of integer elements, non-singular (mod p), such that $C'AC \equiv B$ (mod p), where B is a diagonal matrix, no diagonal element being zero (mod p). Then (28′) becomes

$$(29) \qquad x'Bu \equiv 0, \qquad x'Bx \not\equiv 0, \qquad (\text{mod } p),$$

where $u \equiv C^{-1}v$ and $x \equiv C^{-1}w$ (mod p). Hence we have only to solve

$$(30) \quad a_1u_1x_1 + a_2u_2x_2 + a_3u_3x_3 \equiv 0, \qquad a_1x_1^2 + a_2x_2^2 + a_3x_3^2 \not\equiv 0, \qquad (\text{mod } p),$$

where the a_α are prime to p, and at least one u_α is prime to p since $p \nmid (v_1, v_2, v_3)$.

By symmetry we can suppose $p \nmid u_1$, and solve for x_1 from (30) in the form

$$x_1 \equiv dx_2 + ex_3 \qquad (\text{mod } p).$$

Substituting this in $a_1x_1^2 + a_2x_2^2 + a_3x_3^2$, we get

$$(a_2 + a_1d^2)x_2^2 + 2dex_2x_3 + (a_3 + a_1e^2)x_3^2,$$

which has a coefficient prime to p, the first if $d \equiv 0$, the third if $e \equiv 0$, the second if $de \not\equiv 0$. This completes the proof of Theorem 1.

6. We shall give a few special cases, making use of Charve's table[12] of positive quaternary quadratic forms of determinant $\leqq 20$. Instances of our theory will undoubtedly be far more numerous among indefinite forms, where one class in the genus is the rule rather than the exception; and among forms in which some cross-product coefficients are odd, which may be attained by extending our results through Brandt's generalization of Hermite's identity.

Starting with the matrix $(a_{\alpha\beta})$ of the ternary quadratic form (2, 2, 2, 1, 1, 1) of determinant 4, we have $Q = t_0^2 + 3t_1^2 + 3t_2^2 + 3t_3^2 - 2t_2t_3 - 2t_3t_1 - 2t_1t_2$, and find for the adjoint,

$$4Q^{(3)} = 4(4x_0^2 + 2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_2x_3 + 2x_3x_1 + 2x_1x_2).$$

Since $Q^{(3)}$ is improperly primitive, and there is only one such form of determinant 16 in Charve's table, it belongs to a genus (and order) of one class. The same is therefore true of Q. Since no generic characters are due to Q and $Q^{(3)}$ ($Q^{(2)}$ however represents only $(8n + 3)$'s), S is equivalent to Q. Since Q represents 1 for two values $t_0, \cdots, t_3$, quaternions v with the multiplication table

$$i_1^2 = i_2^2 = i_3^2 = -3, \qquad i_2i_3 = 1 + 2i_1 + i_2 + i_3 = \overline{i_3i_2}, \qquad \text{etc.}$$

[12] Charve, Comptes Rendus, vol. 96(1883), pp. 773–775.

have exactly two right-divisors (t and $-t$) of odd norm m, where m (>0) is assumed to divide Nv but to have no prime factor dividing all of v_0, v_1, v_2, v_3.

Similarly we find that ρ (the number of right-divisors) is 2, if m is prime to 2Δ, for the quaternions arising from $2x_1^2 - 2x_1x_2 + 2x_2^2 + x_3^2$ (having $\Delta = 3$), or from $2x_1^2 + 2x_2^2 + x_3^2$ (having $\Delta = 4$). But $\rho = 4$ if m is prime to 2λ (and the divisors are $\pm t$ and $\pm i_1 t$) in the cases of $\lambda x_1^2 + x_2^2 + x_3^2$ ($\lambda = 2$ or 3), for which the multiplication table is

$$\begin{gathered} i_1^2 = -1, \qquad i_2^2 = i_3^2 = -\lambda, \qquad i_1 i_2 = i_3 = -i_2 i_1, \\ i_3 i_1 = i_2 = -i_1 i_3, \qquad i_2 i_3 = \lambda i_1 = -i_3 i_2. \end{gathered} \tag{31}$$

For ordinary quaternions ($\lambda = 1$), $\rho = 8$ as stated in the introduction.

If there is more than one class in the genus, as for $\lambda = 5$, when there are at least the two classes represented by

$$x_0^2 + x_1^2 + 5x_2^2 + 5x_3^2, \qquad 2x_0^2 - 2x_0x_1 + 3x_1^2 + 2x_2^2 - 2x_2x_3 + 3x_3^2,$$

we may replace the requirement $Nt = p$, in connection with (22), by $Nt = 2p$ (or hp) and deduce that $2v$ has right-divisors of norm $2p$ (or a like theorem).

We mention one type of application. The general solution of $hm^2 = \sum A_{\alpha\beta} v_\alpha v_\beta$ may be obtained by observing that the pure quaternion $v = i_1v_1 + i_2v_2 + i_3v_3$ has its left-divisors the conjugates of its right-divisors, whence $v = \bar{t}wt$, where w is pure and of norm h, and $Nt = m$. For a given h all solutions w of $Nw = h$ can be written down.

7. B. W. Jones and G. Pall[13] have used the results for $\lambda = 1, 2$, and 3 above in proving certain "automorphisms" among the representations of numbers $8n + 1$ or $24n + 1$ in $\lambda x_1^2 + x_2^2 + x_3^2$. E. Rosenthall and C. Solin employ these results in McGill University theses together with that for $f = 2x_1^2 - 2x_1x_2 + 2x_2^2 + x_3^2$ to obtain an arithmetical proof of Glaisher's[14] result for $4(4n + 1) = 2x_1^2 + x_2^2 + x_3^2$, and of new automorphisms for $24n + 1 = f$, $24n + 1 = 3x_1^2 + 3x_2^2 + x_3^2$, and $2(24n + 1) = 3x_1^2 + x_2^2 + x_3^2$.

McGill University.

[13] B. W. Jones and G. Pall, to appear in Acta Mathematica.
[14] J. W. L. Glaisher, Quart. Jour. Math., vol. 20(1884), p. 84.

CERTAIN DIFFERENTIAL EQUATIONS FOR TCHEBYCHEFF POLYNOMIALS

By H. L. Krall

1. **Introduction.** The classical orthogonal polynomials of Jacobi, Laguerre, and Hermite satisfy a differential equation of the form

$$(l_{22}x^2 + l_{21}x + l_{20})y_n''(x) + (l_{11}x + l_{10})y_n'(x) + l_{00}y_n(x) = \lambda_n y_n(x),$$

where the $\{l_{ij}\}$ are constants and λ_n is a parameter. By repeated iterations of this equation one can obtain other differential equations of higher order which have these orthogonal polynomials as solutions. For example, the Legendre polynomials satisfy

$$(x^2 - 1)y_n''(x) + 2xy_n'(x) = n(n + 1)y_n(x),$$

$$(x^2 - 1)^2 y_n^{\text{iv}}(x) + 8x(x^2 - 1)y_n'''(x) + (14x^2 - 6)y_n''(x) + 4xy_n'(x) = n^2(n + 1)^2 y_n(x).$$

However, all the iterates have a special form, namely, the coefficient of the r-th derivative is a polynomial of degree $\leqq r$.

In this paper we shall look for polynomial solutions, in particular, for orthogonal polynomial solutions, of the general differential equation of this type:[1]

$$L(y) = \sum_{i=0}^{r} \left(\sum_{j=0}^{i} l_{ij}x^j \right) y_n^{(i)}(x) = \lambda_n y_n(x), \tag{1}$$

where

$$\lambda_n = l_{00} + nl_{11} + n(n - 1)l_{22} + \cdots .$$

We also consider an extended definition of orthogonal polynomials which we call a Tchebycheff set.

Definition. *Given a set of real or complex constants $\{c_n\}$ such that*

$$\Delta_{nn} = \begin{vmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_1 & c_2 & c_3 & \cdots & c_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n-1} & c_n & c_{n+1} & \cdots & c_{2n-2} \end{vmatrix} \neq 0 \qquad (n = 1, 2, \cdots), \tag{2}$$

Received May 21, 1938. The author takes this opportunity to thank Professor I. M. Sheffer for his many suggestions.

[1] It is obvious that there would be no loss of generality in assuming that $l_{00} = 0$, for this term can be absorbed in the λ_n.

the polynomials

$$(3)\qquad y_n(x) = \begin{vmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_1 & c_2 & c_3 & \cdots & c_{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n-1} & c_n & c_{n+1} & \cdots & c_{2n-1} \\ 1 & x & x^2 & \cdots & x^n \end{vmatrix}, \qquad y_0(x) = c_0,\ (n = 1, 2, \cdots)$$

form a Tchebycheff set.

The condition $\Delta_{nn} \neq 0$ assures us that the degree of $y_n(x)$ is exactly n. This definition becomes the usual definition of orthogonal polynomials if a weight function $\psi(x)$ and an interval (a, b) exist such that

$$c_n = \int_a^b x^n d\psi(x) \qquad (n = 0, 1, \cdots).$$

However, the existence of a function $\psi(x)$ is not implied here.

Our purpose is to find conditions on the coefficients $\{l_{ij}\}$ in order that the differential equation (1) will have a Tchebycheff set of polynomials as solutions.

2. **Formal relations; the function $\mathfrak{D}_n(t)$.** We use the symbol $L(t, x)$ for the power series

$$L(t, x) = \sum_{n=0}^{\infty} L_n(x)t^n,$$

$$L_n(x) = l_{n0} + l_{n1}x + l_{n2}x^2 + \cdots + l_{nn}x^n \qquad (l_{nk} = 0 \quad \text{if} \quad n > r).$$

Rearranging this series to give a power series in x, we obtain

$$L(t, x) = \sum_{n=0}^{\infty} \mathfrak{L}_n(t)x^n, \qquad \mathfrak{L}_n(t) = l_{nn}t^n + l_{n+1,n}\,t^{n+1} + \cdots.$$

This gives a second operator $\mathfrak{L}$:

$$\mathfrak{L}(u(t)) = \sum_{j=0}^{\infty} \mathfrak{L}_j(t)u^j(t).$$

There exists[2] a unique (except for constant multipliers) set of functions $\{\mathfrak{D}_n(t)\}$ $(n = 0, 1, \cdots)$ such that

$$\mathfrak{L}(\mathfrak{D}_n(t)) = \lambda_n \mathfrak{D}_n(t)$$

or

$$(4)\qquad \sum_{i=0}^{r}\sum_{j=0}^{i} l_{ij}t^i \mathfrak{D}_n^j(t) = \lambda_n \mathfrak{D}_n(t),$$

where $\mathfrak{D}_n(t)$ is a formal power series having no powers of t of degree less than n.

[2] I. M. Sheffer, *On the properties of polynomials satisfying a linear differential equation*, Part I, Transactions of the American Mathematical Society, vol. 35(1933), pp. 184–214.

Using a result of Sheffer's, we can obtain the relation[3]

$$\mathfrak{D}_j(t) \cong \sum_{p=0}^{j} \Delta_{jp} K^p(t), \tag{5}$$

where

$$K(t) \cong \sum_{m=0}^{\infty} \frac{c_m}{m!} t^m; \qquad K^p(t) \equiv \frac{d^pK}{dt^p} \cong \sum_{m=0}^{\infty} \frac{c_{m+p}}{m!} t^m, \tag{6}$$

and

$$\Delta_{ni} = \begin{vmatrix} c_0 & c_1 & \cdots & c_{i-1} & c_{i+1} & \cdots & c_n \\ c_1 & c_2 & \cdots & c_i & c_{i+2} & \cdots & c_{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n-1} & c_n & \cdots & c_{n+i-2} & c_{n+i} & \cdots & c_{2n-1} \end{vmatrix}, \tag{7}$$

$$\Delta_{00} \equiv 1, \qquad (n = 1, 2, \cdots; i = 0, 1, \cdots, n-1).$$

To prove relation (5) we use Sheffer's relation[4]

$$e^{tx} \cong \sum_{n=0}^{\infty} \frac{y_n(x)\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \cong \sum_{n=0}^{\infty} \frac{\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \sum_{m=0}^{n} \Delta_{nm} x^m \cong \sum_{m=0}^{\infty} x^m \sum_{n=m}^{\infty} \frac{\mathfrak{D}_n(t)\Delta_{nm}}{\Delta_{nn}\Delta_{n+1,n+1}}.$$

Equating powers of x, we have

$$\frac{t^m}{m!} \cong \sum_{n=m}^{\infty} \frac{\Delta_{nm}\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}}.$$

Then

$$\sum_{p=0}^{j} \Delta_{jp} K^p(t) \cong \sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{\infty} \frac{c_{m+p}t^m}{m!} \cong \sum_{p=0}^{j} \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \frac{c_{m+p}\Delta_{nm}\Delta_{jp}\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}}$$

$$\cong \sum_{n=0}^{\infty} \frac{\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \sum_{p=0}^{j} \sum_{m=0}^{n} c_{m+p}\Delta_{jp}\Delta_{nm}.$$

But Δ_{ni} is the cofactor of c_{n+i} in the last row of $\Delta_{n+1,n+1}$; consequently,

$$\sum_{m=0}^{n} c_{m+p}\Delta_{nm} = \begin{cases} 0, & p < n, \\ \Delta_{n+1,n+1}, & p = n. \end{cases}$$

[3] If $l_{rr} \neq 0$, the formal series, used here, all converge (see Sheffer). If $l_{rr} = 0$, the series, even though they may be formal, may be used, since we equate powers of t or x and only a finite number of terms are necessary to determine the coefficient of any power of t or x. For example, the relation $e^{tx} \cong \sum_{m=0}^{\infty} y_n(x)\mathfrak{D}_n(t)\Delta_{nn}^{-1}\Delta_{n+1,n+1}^{-1}$ states that the coefficients of x^it^j are equal. Only j terms are necessary to find the coefficient on the right side while the coefficient on the left side is zero if $i \neq j$ and $(i!)^{-1}$ if $i = j$. We use $\cong$ to denote the equality of two formal series.

[4] The $\mathfrak{D}_n(t)$ used here is a multiple of the $\mathfrak{D}_n(t)$ used by Sheffer. In fact, here $\mathfrak{D}_n(t) = \Delta_{n+1,n+1}t^n + \cdots$, while in Sheffer's paper $\mathfrak{D}_n(t) = t^n + \cdots$.

Thus

$$\sum_{p=0}^{j}\sum_{m=0}^{n} c_{m+p}\Delta_{jp}\Delta_{nm} = \begin{cases} \sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{n} c_{m+p}\Delta_{nm} = 0, & n > j, \\ \sum_{m=0}^{n} \Delta_{nm} \sum_{p=0}^{j} c_{m+p}\Delta_{jp} = 0, & j > n, \\ \Delta_{jj}\Delta_{j+1,j+1}, & n = j. \end{cases}$$

And finally

$$\sum_{p=0}^{j} \Delta_{jp} K^p(t) \cong \mathfrak{D}_j(t).$$

Substituting the relation for $\mathfrak{D}_j(t)$ in (5) into (4), we get

$$\sum_{i=0}^{r}\sum_{j=0}^{i}\sum_{l=0}^{n} l_{ij} t^i \Delta_{nl} K^{l+j}(t) \cong \lambda_n \sum_{l=0}^{n} \Delta_{nl} K^l(t).$$

Using the expansion for $K^p(t)$, we get

$$\sum_{i=0}^{r}\sum_{j=0}^{i}\sum_{l=0}^{n}\sum_{m=0}^{\infty} \frac{c_{m+l+j}}{m!} l_{ij}\Delta_{nl} t^i t^m \cong \lambda_n \sum_{l=0}^{n}\sum_{m=0}^{\infty} \Delta_{nl}\frac{c_{m+l}}{m!} t^m.$$

If we replace m by $m - i$, this becomes

$$\sum_{i=0}^{r}\sum_{j=0}^{i}\sum_{l=0}^{n}\sum_{m=i}^{\infty} \frac{c_{m-i+l+j}}{(m-i)!} l_{ij}\Delta_{nl} t^m \cong \lambda_n \sum_{l=0}^{n}\sum_{m=0}^{\infty} \Delta_{nl}\frac{c_{m+l}}{m!} t^m.$$

Equating the coefficients of t^m and agreeing to drop terms containing negative factorials, we have

$$\sum_{i=0}^{r}\sum_{j=0}^{i}\sum_{l=0}^{n} l_{ij}\Delta_{nl}\frac{c_{m-i+l+j}}{(m-i)!} = \lambda_n \sum_{l=0}^{n} \Delta_{nl}\frac{c_{m+l}}{m!}.$$

Let $j = i - u$, then

$$\sum_{i=0}^{r}\sum_{u=0}^{i}\sum_{l=0}^{n} l_{i,i-u}\Delta_{nl}\frac{c_{m+l-u}}{(m-i)!} = \lambda_n \sum_{l=0}^{u} \Delta_{nl}\frac{c_{m+l}}{m!}.$$

Multiply both sides by $m!$ and use the notation

(8) $$x^{(n)} = x(x-1)\cdots(x-n+1); \qquad x^{(0)} = 1.$$

We have, assuming (as we may) that $l_{00} = 0$,

(9) $$B_n(m) = \sum_{l=0}^{n} \Delta_{nl} \sum_{i=1}^{r}\sum_{u=0}^{i} m^{(i)} l_{i,i-u} c_{m+l-u} - \lambda_n \sum_{l=0}^{n} \Delta_{nl} c_{m+l} = 0.$$

The first relation $B_0(m) = 0$ is sufficient to determine the $\{c_n\}$ in terms of c_0, hence the other recurrence relations are extra conditions on the $\{l_{ij}\}$. These conditions $B_n(m) = 0$, together with $\Delta_{nn} \neq 0$ $(n = 0, 1, \cdots)$, are necessary and sufficient for the existence of a Tchebycheff set satisfying the differential equa-

tion. To prove the sufficiency we assume (i) $B_n(m) = 0$, (ii) $\Delta_{nn} \neq 0$ ($n = 0, 1, \cdots$) and proceed as follows.

Set

$$\frac{t^m}{m!} \cong \sum_{n=m}^{\infty} \frac{\Delta_{nm}\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} + Q_{m+1}(t),$$

where

$$Q_{m+1}(t) \cong \sum_{k=1}^{\infty} q_{m+1,m+k} t^{m+k}.$$

Then

$$\mathfrak{D}_j(t) \cong \sum_{p=0}^{j} \Delta_{jp} K^p(t) \cong \sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{\infty} c_{m+p} \frac{t^m}{m!}$$

$$\cong \sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{\infty} c_{m+p} \left\{ \sum_{n=m}^{\infty} \frac{\Delta_{nm}\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} + Q_{m+1}(t) \right\}.$$

Also

$$\mathfrak{D}_j(t) \cong \sum_{n=0}^{\infty} \frac{\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \sum_{p=0}^{j} \sum_{m=0}^{n} c_{m+p}\Delta_{jp}\Delta_{nm}$$

$$\cong \sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{\infty} c_{m+p} \sum_{n=m}^{\infty} \frac{\Delta_{nm}\mathfrak{D}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}}.$$

These two relations give us

$$\sum_{p=0}^{j} \Delta_{jp} \sum_{m=0}^{\infty} c_{m+p} Q_{m+1}(t) \cong 0, \qquad \sum_{m=0}^{\infty} Q_{m+1}(t) \sum_{p=0}^{j} \Delta_{jp} c_{m+p} \cong 0.$$

We get for

$$j = 0 : \Delta_{00} \sum_{m=0}^{\infty} c_m Q_{m+1}(t) \cong 0,$$

$$j = 1 : \sum_{m=0}^{\infty} Q_{m+1}(t)\{c_m\Delta_{10} + c_{m+1}\Delta_{11}\} \cong \Delta_{11} \sum_{m=0}^{\infty} Q_{m+1}(t)c_{m+1} \cong 0,$$

$$j = 2 : \sum_{m=0}^{\infty} Q_{m+1}(t)\{c_m\Delta_{20} + c_{m+1}\Delta_{21} + c_{m+2}\Delta_{22}\} \cong \Delta_{22} \sum_{m=0}^{\infty} Q_{m+1}(t)c_{m+2} \cong 0.$$

In general, since $\Delta_{nn} \neq 0$,

$$\sum_{m=0}^{\infty} Q_{m+1}(t)c_{m+s} \cong 0,$$

$$0 \cong \sum_{m=0}^{\infty} c_{m+s} \sum_{k=1}^{\infty} q_{m+1,m+k} t^{m+k} \cong \sum_{m=0}^{\infty} c_{m+s} \sum_{k=m}^{\infty} q_{m+1,k+1} t^{k+1},$$

$$0 \cong \sum_{k=0}^{\infty} t^{k+1} \sum_{m=0}^{k} q_{m+1,k+1} c_{m+s}.$$

Thus,

$$\sum_{m=0}^{k} q_{m+1,k+1}c_{m+s} = 0.$$

Since $\Delta_{nn} \neq 0$, we have for

$$k = 0, s = 0 : q_{11}c_0 = 0, q_{11} = 0;$$

$$k = 1, s = 0 : q_{12}c_0 + q_{22}c_1 = 0;$$

$$k = 1, s = 1 : q_{12}c_1 + q_{22}c_2 = 0, q_{12} = q_{22} = 0.$$

It follows at once that $Q_{m+1}(t) = 0$, and then

$$\frac{t^m}{m!} \cong \sum_{n=m}^{\infty} \frac{\Delta_{nm}\mathfrak{T}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}},$$

$$\sum_{n=0}^{\infty} \frac{y_n(x)\mathfrak{T}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \cong \sum_{n=0}^{\infty} \frac{\mathfrak{T}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \sum_{m=0}^{n} \Delta_{nm} x^m$$

$$\cong \sum_{m=0}^{\infty} x^m \sum_{n=m}^{\infty} \frac{\mathfrak{T}_n(t)\Delta_{nm}}{\Delta_{nn}\Delta_{n+1,n+1}} \cong \sum_{m=0}^{\infty} \frac{x^m t^m}{m!} \cong e^{tx}.$$

Since $B_n(m) = 0$, the function $\mathfrak{T}_n(t)$ must satisfy (4). Sheffer has shown that $\mathfrak{L}(e^{tx}) = L(e^{tx})$. Accordingly,

$$\sum_{n=0}^{\infty} \frac{L(y_n(x))\mathfrak{T}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}} \cong L(e^{tx}) \cong \mathfrak{L}(e^{tx}) \cong \sum_{n=0}^{\infty} \frac{y_n(x)\mathfrak{L}(\mathfrak{T}_n(t))}{\Delta_{nn}\Delta_{n+1,n+1}}$$

$$\cong \sum_{n=0}^{\infty} \frac{\lambda_n y_n(x)\mathfrak{T}_n(t)}{\Delta_{nn}\Delta_{n+1,n+1}}.$$

Equating the coefficients of $\mathfrak{T}_n(t)$, we get finally

$$L(y_n(x)) = \lambda_n y_n(x).$$

If we find a set of $\{c_n\}$ and $\{l_{ij}\}$ which satisfy $B_n(m) = 0$ and $\Delta_{nn} \neq 0$, we can immediately write out the differential equation and the solutions $\{y_n(x)\}$. The classical polynomials furnish us with examples of equations of any even order; however, the question of the existence of equations whose solutions form a non-classical Tchebycheff set is more difficult. In order to answer this question we proceed to replace (9) by equivalent recurrence relations which are much simpler.

3. **Some lemmas.** We shall use the following well known formulas:

$$\Delta[u(m)\cdot v(m)] = u(m+1)v(m+1) - u(m)v(m)$$

$$= u(m+1)\Delta v(m) + v(m)\Delta u(m),$$

$$\Delta(m+\alpha)^{(k)} = k(m+\alpha)^{(k-1)}.$$

LEMMA 1.

$$m^{(2i)} - l^{(2i)} = (m - l) \sum_{k=0}^{i-1} (-1)^k \binom{2i - k - 1}{k} (m + l - 2k - 1)^{(2i-2k-1)} m^{(k)} l^{(k)}.$$

Proof. We consider the right side, R,

$$\begin{aligned}\Delta_m R = (m + 1 - l) \sum_{k=0}^{i-1} (-1)^k \binom{2i - k - 1}{k} \{&(m + l - 2k)^{(2i-2k-1)} k m^{(k-1)} l^{(k)} \\ &+ (2i - 2k - 1)(m + l - 2k - 1)^{(2i-2k-2)} m^{(k)} l^{(k)}\} \\ &+ \sum_{k=0}^{i-1} (-1)^k \binom{2i - k - 1}{k} \{(m + l - 2k - 1)^{(2i-2k-1)} m^{(k)} l^{(k)}\},\end{aligned}$$

$$\begin{aligned}\Delta_l \Delta_m R = (m - l) \sum_{k=0}^{i-1} (-1)^k \binom{2i - k - 1}{k} \{&(m + l - 2k + 1)^{(2i-2k-1)} k^2 m^{(k-1)} l^{(k-1)} \\ &+ (2i - 2k - 1)(m + l - 2k)^{(2i-2k-2)} k m^{(k-1)} l^{(k)} \\ &+ (2i - 2k - 1)(m + l - 2k)^{(2i-2k-2)} m^{(k)} k l^{(k-1)} \\ &+ (2i - 2k - 2)(2i - 2k - 1)(m + l - 2k - 1)^{(2i-2k-3)} m^{(k)} l^{(k)}\} \\ &+ \sum_{k=0}^{i-1} (-1)^k \binom{2i - k - 1}{k} \{-(m + l - 2k)^{(2i-2k-1)} k m^{(k-1)} l^{(k)} \\ &- (2i - 2k - 1)(m + l - 2k - 1)^{(2i-2k-2)} m^{(k)} l^{(k)} \\ &+ (m + l - 2k)^{(2i-2k-1)} m^{(k)} k l^{(k-1)} \\ &+ (2i - 2k - 1)(m + l - 2k - 1)^{(2i-2k-2)} m^{(k)} l^{(k)}\}.\end{aligned}$$

These eight terms can be simplified as follows. The sixth and eighth terms cancel. The second and third terms give

$$\begin{aligned}(2i - 2k - 1)&(m + l - 2k)^{(2i-2k-2)} k m^{(k-1)} l^{(k-1)} (m + l - 2k + 2) \\ &= k(2i - 2k - 1)(m + l - 2k + 1)^{(2i-2k-1)} m^{(k-1)} l^{(k-1)} \\ &\quad + k(2i - 2k - 1)(m + l - 2k)^{(2i-2k-2)} m^{(k-1)} l^{(k-1)}.\end{aligned}$$

The fifth and seventh terms give

$$\begin{aligned}k(m + l - 2k)^{(2i-2k-1)} &\{m^{(k)} l^{(k-1)} - m^{(k-1)} l^{(k)}\} \\ &= (m - l) k (m + l - 2k)^{(2i-2k-1)} m^{(k-1)} l^{(k-1)} \\ &= (m - l) k m^{(k-1)} l^{(k-1)} \{(m + l - 2k + 1)^{(2i-2k-1)} \\ &\qquad - (2i - 2k - 1)(m + l - 2k)^{(2i-2k-2)}\}.\end{aligned}$$

Thus, since two more terms cancel,

$$\begin{aligned}\Delta_l\Delta_m R &= (m-l)\sum_{k=0}^{i-1}(-1)^k\binom{2i-k-1}{k}\\
&\qquad\{(m+l-2k+1)^{(2i-2k-1)}m^{(k-1)}l^{(k-1)}(k^2+k(2i-2k-1)+k)\\
&\qquad+(2i-2k-1)(2i-2k-2)(m+l-2k-1)^{(2i-2k-3)}m^{(k)}l^{(k)}\}\\
&= (m-l)\sum_{k=0}^{i-1}(-1)^k\binom{2i-k-1}{k}\\
&\qquad\{k(2i-k)(m+l-2k+1)^{(2i-2k-1)}m^{(k-1)}l^{(k-1)}\}\\
&\quad+(m-l)\sum_{k=0}^{i-1}(-1)^k\binom{2i-k-1}{k}\\
&\qquad\{(2i-2k-1)(2i-2k-2)(m+l-2k-1)^{(2i-2k-3)}m^{(k)}l^{(k)}\}.\end{aligned}$$

If we replace k by $k+1$ in the first summation, we get the negative of the second. Accordingly, $\Delta_l\Delta_m R = 0$. Since R is a polynomial in m and l, it must be of the form

$$R = f(m) - g(l).$$

Setting $m = 0$ we have $g(l) = l^{(2i)}$; setting $l = 0$ we have $f(m) = m^{(2i)}$. This proves the lemma.

LEMMA 2.

$$m^{(2i-1)} - l^{(2i-1)} = (m-l)\sum_{k=0}^{i-1}(-1)^k\binom{2i-k-2}{k}(m+l-2k-1)^{(2i-2k-2)}m^{(k)}l^{(k)}.$$

We omit the proof of this lemma, for it is almost identical with the proof of Lemma 1.

LEMMA 3.

$$m^{(i)} - (m-l)\sum_{k=0}^{n-1}(-1)^k\binom{i-k-1}{k}(m+l-2k-1)^{(i-2k-1)}m^{(k)}l^{(k)}$$

$$= \begin{cases}0, & \textit{when}\quad l = 0, 1, \cdots, n-1,\\ (-1)^n n!\dbinom{i-n-1}{n}m^{(i-n)}, & \textit{when}\quad l = n,\end{cases}$$

where $i > 2n > 0$.

Proof. If i is even, replace i by $2i$ and the left side becomes

$$\begin{aligned}L &= m^{(2i)} - (m-l)\sum_{k=0}^{n-1}(-1)^k\binom{2i-k-1}{k}(m+l-2k-1)^{(2i-2k-1)}m^{(k)}l^{(k)}\\
&= l^{(2i)} + (m-l)\sum_{k=n}^{i-1}(-1)^k\binom{2i-k-1}{k}(m+l-2k-1)^{(2i-2k-1)}m^{(k)}l^{(k)}.\end{aligned}$$

Since $l^{(k)} = 0$ when $l = 0, 1, \cdots, k-1$, we have $L = 0$ when $l = 0, 1, \cdots, n-1$. When $l = n$, we have

$$L = (m-n)(-1)^n \binom{2i-n-1}{n} (m+n-2n-1)^{(2i-2n-1)} m^{(n)} l^{(n)}$$
$$= (-1)^n n! \binom{2i-n-1}{n} m^{(2i-n)}.$$

If i is odd, replace i by $2i-1$, and we get

$$L = m^{(2i-1)} - (m-l) \sum_{k=0}^{n-1} (-1)^k \binom{2i-k-2}{k} (m+l-2k-1)^{(2i-2k-2)} m^{(k)} l^{(k)}$$
$$= l^{(2i-1)} + (m-l) \sum_{k=n}^{i-1} (-1)^k \binom{2i-k-2}{k} (m+l-2k-1)^{(2i-2k-2)} m^{(k)} l^{(k)}.$$

As before, $L = 0$ when $l = 0, 1, \cdots, n-1$, and when $l = n$,

$$L = (m-n)(-1)^n \binom{2i-n-2}{n} (m+n-2n-1)^{(2i-2n-2)} m^{(n)} l^{(n)}$$
$$= (-1)^n n! \binom{2i-n-2}{n} m^{(2i-n-1)}.$$

4. **Fundamental recurrence relations.** Let

$$S_n(m) \tag{10}$$
$$= \begin{cases} \displaystyle\sum_{i=2n+1}^{r} \sum_{u=0}^{i} \binom{i-n-1}{n} (m-2n-1)^{(i-2n-1)} l_{i,i-u} c_{m-u}, & \text{when } 2n+1 \leqq r, \\ 0, & \text{when } 2n+1 > r. \end{cases}$$

Then (see (9)), we have

LEMMA 4.

$$\begin{aligned} B_n(m) - \sum_{l=0}^{n} \Delta_{nl}(m-l) \sum_{k=0}^{n-1} (-1)^k m^{(k)} l^{(k)} S_k(m+l) \\ = (-1)^n n! \Delta_{nn} m^{(n+1)} S_n(m+n) + \sum_{l=0}^{n-1} K(n,l) c_{m+l} \end{aligned} \tag{11}$$

where

$$K(n,l) = \sum_{u=1}^{n-l} \sum_{i=u}^{l+u} \Delta_{n,l+u} (l+u)^{(i)} l_{i,i-u} + \sum_{i=1}^{n} \Delta_{nl} (l^{(i)} - n^{(i)}) l_{ii}.$$

The particular form of $K(n, l)$ is unimportant for our purpose. However, it is important that $K(n, l)$ does not depend on m and that the last summation on the right of (11) extends from $l = 0$ to $l = n - 1$.

Proof. First assume $2n + 1 \leqq r$. Then let

$$Q = \sum_{k=0}^{n-1} (-1)^k m^{(k)} l^{(k)} S_k(m + l) = \sum_{k=0}^{n-1}\left(\sum_{i=2n+1}^{r} + \sum_{i=2k+1}^{2n-1}{}' + \sum_{i=2k+2}^{2n}{}' \left\{ \sum_{u=0}^{i} H(i) \right\}\right),$$

where

$$H(i) = (-1)^k \binom{i-k-1}{k} m^{(k)} l^{(k)} (m + l - 2k - 1)^{(i-2k-1)} l_{i,i-u} c_{m+l-u}$$

and $\sum'$ denotes summation through odd (or even) values of i.

$$Q = \sum_{k=0}^{n-1} \sum_{i=2n+1}^{r} \sum_{u=0}^{i} H(i) + \sum_{k=0}^{n-1} \sum_{i=k}^{n-1} \sum_{u=0}^{2i+1} H(2i + 1) + \sum_{k=0}^{n-1} \sum_{i=k}^{n-1} \sum_{u=0}^{2i+2} H(2i + 2)$$

$$= \sum_{i=2n+1}^{r} \sum_{u=0}^{i} \sum_{k=0}^{n-1} H(i) + \sum_{i=0}^{n-1} \sum_{k=0}^{i} \sum_{u=0}^{2i+1} H(2i + 1) + \sum_{i=0}^{n-1} \sum_{k=0}^{i} \sum_{u=0}^{2i+2} H(2i + 2),$$

$$Q = \sum_{i=2n+1}^{r} \sum_{u=0}^{i} \sum_{k=0}^{n-1} H(i) + \sum_{i=1}^{n} \sum_{u=0}^{2i-1} \sum_{k=0}^{i-1} H(2i - 1) + \sum_{i=1}^{n} \sum_{u=0}^{2i} \sum_{k=0}^{i-1} H(2i). \tag{12}$$

Also from (9)

$$B_n(m) = \sum_{l=0}^{n} \Delta_{nl} \left[\left(\sum_{i=2n+1}^{r} + \sum_{i=1}^{2n-1}{}' + \sum_{i=2}^{2n}{}' \right) \left(\sum_{u=0}^{i} m^{(i)} l_{i,i-u} c_{m+l-u} \right) \right]$$
$$- \sum_{l=0}^{n} \Delta_{nl} \sum_{i=1}^{n} n^{(i)} l_{ii} c_{m+l}.$$

If we set $F(i) = m^{(i)} l_{i,i-u} c_{m+l-u}$, this becomes

$$\begin{aligned} B_n(m) = \sum_{l=0}^{n} \Delta_{nl} \Bigg[& \sum_{i=2n+1}^{r} \sum_{u=0}^{i} F(i) + \sum_{i=1}^{n} \sum_{u=0}^{2i-1} F(2i - 1) \\ & + \sum_{i=1}^{n} \sum_{u=0}^{2i} F(2i) - \sum_{i=1}^{n} n^{(i)} l_{ii} c_{m+l} \Bigg]. \end{aligned} \tag{13}$$

If we use (12) and (13), the left side, L, of (11) becomes

$$\begin{aligned} L = \sum_{l=0}^{n} \Delta_{nl} \Bigg[& \sum_{i=2n+1}^{r} \sum_{u=0}^{i} \left\{ F(i) - (m - l) \sum_{k=0}^{n-1} H(i) \right\} \\ & + \sum_{i=1}^{n} \sum_{u=0}^{2i-1} \left\{ F(2i - 1) - (m - l) \sum_{k=0}^{i-1} H(2i - 1) \right\} \\ & + \sum_{i=1}^{n} \sum_{u=0}^{2i} \left\{ F(2i) - \sum_{k=0}^{i-1} H(2i) \right\} - \sum_{i=1}^{n} n^{(i)} l_{ii} c_{m+l} \Bigg]. \end{aligned}$$

Lemmas 3, 2 and 1 apply to the terms in braces; hence,

$$L = \Delta_{nn} \sum_{i=2n+1}^{r} \sum_{u=0}^{i} (-1)^n n! \binom{i-n-1}{n} m^{(i-n)} l_{i,i-u} c_{m+n-u}$$

$$+ \sum_{l=0}^{n} \Delta_{nl} \left[\sum_{i=1}^{n} \sum_{u=0}^{2i-1} l^{(2i-1)} l_{2i-1,2i-1-u} c_{m+l-u} \right.$$

$$\left. + \sum_{i=1}^{n} \sum_{u=0}^{2i} l^{(2i)} l_{2i,2i-u} c_{m+l-u} - \sum_{i=1}^{n} n^{(i)} l_{ii} c_{m+l} \right]$$

$$= (-1)^n \Delta_{nn} n!\, m^{(n+1)} S_n(m+n)$$

$$+ \sum_{l=0}^{n} \Delta_{nl} \left[\sum_{i=1}^{2n} \sum_{u=0}^{i} l^{(i)} l_{i,i-u} c_{m+l-u} - \sum_{i=1}^{n} n^{(i)} l_{ii} c_{m+l} \right].$$

We separate the term for which $u = 0$, use the fact that $l^{(i)} = 0$ when $i > l$, and get

$$L = (-1)^n \Delta_{nn} n!\, m^{(n+1)} S_n(m+n) + \sum_{l=1}^{n} \sum_{i=1}^{l} \sum_{u=1}^{i} \Delta_{nl} l^{(i)} l_{i,i-u} c_{m+l-u}$$

$$+ \sum_{l=0}^{n} \Delta_{nl} \sum_{i=1}^{n} (l^{(i)} - n^{(i)}) l_{ii} c_{m+l}.$$

Since $u \leqq i \leqq l$, the first sum contains no c_n whose subscript is less than m; also the greatest subscript in this sum is $m + n - 1$. In the second sum, the coefficient of c_{m+n} is zero, so rearranging the terms we have

$$L = (-1)^n \Delta_{nn} n! m^{(n+1)} S_n(m+n) + \sum_{l=0}^{n-1} K(n, l) c_{m+l}.$$

To obtain the exact form of $K(n, l)$, rearrange the second term as follows:

$$\sum_{l=1}^{n} \sum_{i=1}^{l} \sum_{u=1}^{i} \Delta_{nl} l^{(i)} l_{i,i-u} c_{m+l-u} = \sum_{u=1}^{n} \sum_{l=u}^{n} \sum_{i=u}^{l} \Delta_{nl} l^{(i)} l_{i,i-u} c_{m+l-u}.$$

Replace l by $l + u$; this becomes

$$\sum_{u=1}^{n} \sum_{l=0}^{n-u} \sum_{i=u}^{l+u} \Delta_{n,l+u} (l+u)^{(i)} l_{i,i-u} c_{m+l} = \sum_{l=0}^{n-1} \sum_{u=1}^{n-l} \sum_{i=u}^{l+u} \Delta_{n,l+u} (l+u)^{(i)} c_{m+l}$$

and

$$K(n, l) = \sum_{u=1}^{n-l} \sum_{i=u}^{l+u} \Delta_{n,l+u} (l+u)^{(i)} l_{i,i-u} + \Delta_{nl} \sum_{i=1}^{n} \{l^{(i)} - n^{(i)}\} l_{ii}.$$

Now assume $2n + 1 > r$. We define $l_{i,i-u} = 0$ for $i > r$, and the proof is the same as the above. It also follows from this definition that $S_n(m) = 0$ for $2n + 1 > r$.

THEOREM 1. *If* $\Delta_{nn} \neq 0$ $(n = 0, 1, \cdots)$, *then, in order that*

(14) $$B_n(m) = 0 \qquad (m, n = 0, 1, \cdots),$$

it is necessary and sufficient that

$$S_n(m) = 0 \qquad (n = 0, 1, \cdots; m = 2n + 1, 2n + 2, \cdots). \tag{15}$$

Proof. First assume (14). We shall prove (15) by induction. We have from (11)

$$mS_0(m) = \sum_{i=1}^{r} \sum_{u=0}^{i} m^{(i)} l_{i,i-u} c_{m-u} = B_0(m).$$

Hence $S_0(m) = 0$ for $m = 1, 2, \cdots$. Assume $S_k(m) = 0$ $(k = 0, 1, \cdots, n - 1; m = 2k + 1, 2k + 2, \cdots)$. We have (from Lemma 4)

$$\begin{aligned} -\sum_{l=0}^{n} \Delta_{nl}(m - l) \sum_{k=0}^{n-1} (-1)^k m^{(k)} l^{(k)} S_k(m + l) \\ = (-1)^n \Delta_{nn} n! m^{(n+1)} S_n(m + n) + \sum_{l=0}^{n-1} K(n, l) c_{m+l}. \end{aligned} \tag{16}$$

Now $(m - l)m^{(k)}l^{(k)}S_k(m + l)$ is zero if $m < k$ because of the factor $m^{(k)}$; it is zero if $l < k$ because of the factor $l^{(k)}$; it is zero if $m = l = k$ because of the factor $m - l$; it is zero if $m + l > 2k$ because of the factor $S_k(m + l)$. (16) reduces to

$$(-1)^n \Delta_{nn} m^{(n+1)} S_n(m + n) + \sum_{l=0}^{n-1} K(n, l) c_{m+l} = 0.$$

Substitute $m = 0, 1, \cdots, n - 1$ in this relation. The first term is zero because of the factor $m^{(n+1)}$, and we get n equations in the n unknowns $K(n, l)$ $(l = 0, 1, \cdots, n - 1)$. Since the determinant of the coefficients, Δ_{nn}, is not zero, all the K's must be zero and

$$S_n(m + n) = 0 \qquad (m = n + 1, n + 2, \cdots)$$

or

$$S_n(m) = 0 \qquad (m = 2n + 1, 2n + 2, \cdots).$$

Now assume (15). Then (11) becomes

$$B_n(m) = \sum_{l=0}^{n-1} K(n, l) c_{m+l}.$$

The function $B_n(m)$ can be written in the form

$$B_n(m) = \sum_{i=1}^{r} \sum_{u=0}^{i} m^{(i)} l_{i,i-u} \left[\sum_{l=0}^{n} \Delta_{nl} c_{m+l-u} \right] - \lambda_n \sum_{l=0}^{n} \Delta_{nl} c_{m+l}.$$

When $m = 0, 1, \cdots, n - 1$, the term in brackets is the summation of the elements of one row of $\Delta_{n+1,n+1}$ multiplied by the co-factors of another. Hence

$$\sum_{l=0}^{n-1} K(n, l) c_{m+l} = 0 \qquad (m = 0, 1, \cdots, n - 1).$$

Since $\Delta_{nn} \neq 0$, we have $K(n, l) \equiv 0$, and

$$B_n(m) = 0 \qquad (m = 0, 1, \cdots).$$

THEOREM 2. *In order that there exist a Tchebycheff set $\{y_n(x)\}$, satisfying the differential equation*

$$\sum_{i=0}^{r} (l_{i0} + l_{i1}x + l_{i2}x^2 + \cdots + l_{ii}x^i)y_n^{(i)}(x) = \lambda_n y_n(x), \tag{17}$$

where $\lambda_n = l_{00} + nl_{11} + n(n-1)l_{22} + \cdots$, it is necessary and sufficient that (i) *the moments satisfy the recurrence relations*

$$S_n(m) = \sum_{i=2n+1}^{r} \sum_{u=0}^{i} \binom{i-n-1}{n} (m-2n-1)^{(i-2n-1)} l_{i,i-u} c_{m-u} = 0$$

$$(2n + 1 \leqq r;\ m = 2n + 1, 2n + 2, \cdots),$$

and (ii) $\Delta_{nn} \neq 0$ $(n = 0, 1, \cdots)$.

This theorem follows at once from Theorem 1 and from the fact that $\{y_n(x)\}$ satisfy (1) if and only if the conditions $B_n(m) = 0$; $\Delta_{nn} \neq 0$ $(n = 0, 1, \cdots)$ are fulfilled.

5. **Application of Theorem 2.** Let us first consider the case when r is odd. The recurrence relation with $2n + 1 = r$ is

$$l_{rr}c_m + l_{r,r-1}c_{m-1} + \cdots + l_{r0}c_{m-r} = 0.$$

If we multiply the first row of (3) by l_{r0}, the second row by l_{r1}, $\cdots$ and add to the r-th row, we get a row of zeros. Accordingly, $y_n(x) \equiv 0$ when $n > r$.

THEOREM 3. *There is no differential equation of type* (17) *of odd order having a Tchebycheff set as solutions.*

However, if r is even, we do get something new, i.e., there are differential equations of even order with non-classical orthogonal polynomial solutions. If we look for a non-classical fourth-order differential equation with $0 = c_1 = c_3 = \cdots$, we find the example

$$(x^2 - 1)^2 y_n^{\text{iv}}(x) + 8x(x^2 - 1)y_n'''(x) + (4\alpha + 12)(x^2 - 1)y_n''(x) + 8\alpha x y_n'(x)$$
$$= n(n + 1)(n^2 + n + 4\alpha - 2)y_n(x),$$

where α is an arbitrary parameter. The recurrence relations can be reduced to

$$(m + 1)(m + \alpha - 1)c_m - (m - 1)(m + \alpha + 1)c_{m-2} = 0,$$

$$(m + 1)c_m - 2(m - 1)c_{m-2} + (m - 3)c_{m-4} = 0.$$

We find that

$$c_{2m} = \frac{2m + 1 + \alpha}{(1 + \alpha)(2m + 1)} c_0 = \left(1 + \frac{\alpha}{2m + 1}\right) \frac{c_0}{1 + \alpha} = \frac{c_0}{1 + \alpha} \int_{-1}^{+1} x^{2m}\, d\psi(x)$$

$$(c_0 \text{ arbitrary}, n \geqq 1),$$

where

$$\psi(x) = \begin{cases} 1 + \alpha, & \text{when } x = 1, \\ \alpha x, & \text{when } -1 < x < 1, \\ -(1 + \alpha), & \text{when } x = -1. \end{cases}$$

As a particular case, take $\alpha = 1$, then $\psi(x)$ is monotonically increasing, and we know from the classical theory that $\Delta_{nn} > 0$. In this case we have

$$(18) \quad L(y) = (x^2 - 1)^2 y_n^{\text{iv}}(x) + 8x(x^2 - 1)y_n'''(x) + 16(x^2 - 1)y_n''(x) + 8xy_n'(x) = n(n + 1)(n^2 + n + 2)y_n(x).$$

It is easy to verify that the above weight function $\psi(x)$ (with $\alpha = 1$)[5] generates a Tchebycheff set which satisfies (18). To do this we notice that

$$\int_{-1}^{+1} \{vL(u) - uL(v)\}\, d\psi(x) = \{vL(u) - uL(v)\}_{x=-1} + \{vL(u) - uL(v)\}_{x=1} + \{(12x^2 - 4)(uv' - vu')\}_{-1}^{+1} = 0.$$

Let $\{P_n(x)\}$ be the Tchebycheff set generated by $\psi(x)$, and let $G_{n-1}(x)$ be an arbitrary polynomial of degree $n - 1$. Substitute $u = P_n(x)$, $v = G_{n-1}(x)$ in the above relation. Since $L(G_{n-1}(x))$ is also a polynomial of degree $n - 1$, we get

$$0 = \int_{-1}^{+1} \{G_{n-1}(x)L(P_n(x)) - P_n(x)L(G_{n-1}(x))\}\, d\psi(x) = \int_{-1}^{+1} G_{n-1}(x)L(P_n(x))\, d\psi(x).$$

It follows that the n-th degree polynomial, $L(P_n(x))$, must be a multiple of $P_n(x)$, i.e.,

$$L(P_n(x)) = \lambda_n P_n(x).$$

PENNSYLVANIA STATE COLLEGE.

[5] It is interesting to note that this $\psi(x)$ is obtained from the $\psi(x)$ for Legendre polynomials by adding 2 masses ($= 1$) at $x = \pm 1$.

PROPERTIES OF INVARIANT SETS UNDER POINTWISE PERIODIC HOMEOMORPHISMS

By D. W. Hall and G. E. Schweigert

A single-valued continuous transformation $T(M) = M$ of a compact metric space onto itself is said to be *pointwise periodic* provided that for each point x in M there exists a positive integer n such that $T^n(x) = x$. It follows directly from this definition that such a transformation is one-to-one and hence is a homeomorphism of M onto itself. W. L. Ayres has recently studied this type of homeomorphism together with other related types on locally connected continua.[1] His generosity in discussing these results has interested the authors in this type of transformation and thus led to the present paper.

In contrast with previous papers we require only that the space be compact and metric.

1. **Definitions and preliminary lemmas**. Let M be a compact metric space and $T(M) = M$ a pointwise periodic homeomorphism. Let L (or L_0) be any closed invariant subset of M, i.e., any closed subset of M such that $T(L) = L$. Denote by $p(T, x)$ the period of any point x in M under T, i.e., the least positive integer n such that $T^n(x) = x$. Let L_1 consist of all those points x of L such that for any positive integer N and any neighborhood U of x there exists a point y (distinct from x) in U for which $p(T, y) > N$. In other words, L_1 consists of all points of L at which $p(T, x)$ has an unbounded limit superior. If L_β has been defined for all ordinals β less than a given ordinal α, we may define L_α as follows. In case α is an isolated number, let L_α consist of all those points x of $L_{\alpha-1}$ at which $p(T, x)$ has an unbounded limit superior (T is considered as being defined only on $L_{\alpha-1}$). If α is a limit number, define $L_\alpha = \prod_{\beta<\alpha} L_\beta$. Thus we have defined for every ordinal α a set L_α.

Lemma 1. *For every α the set L_α is closed and invariant.*

Proof. The proof will be by transfinite induction. Assume the lemma true for all ordinals $\beta < \alpha$. Then if α is a limit number, L_α is closed, being the product of closed sets. It is also invariant, since if $p \in L_\alpha$ then $p \in L_\beta$ for all $\beta < \alpha$. Thus $T(p) \in L_\beta$ for all $\beta < \alpha$ since all these sets are invariant by hypothesis. Consequently, $T(p) \in L_\alpha$ so that this set is invariant.

On the other hand, if α is an isolated number, then $L_{\alpha-1}$ exists. For any point $p \in L_\alpha$ we may find a sequence of points $\{p_i\}$ in $L_{\alpha-1}$ such that $\lim p_i = p$

Received August 12, 1938. The first-mentioned author is a National Research Fellow

[1] See Bulletin of the American Mathematical Society, vol. 44(1938), p. 329, abstract no. 172; Bulletin of the American Mathematical Society, vol. 43(1937), p. 20, abstract no. 3; and a forthcoming article in Fundamenta Mathematicae.

and $p(T, p_i)$ is monotone increasing with i. By hypothesis $L_{\alpha-1}$ is closed and invariant, so $T(p_i) \epsilon L_{\alpha-1}$ for all i and $\lim T(p_i) = T(p)$, some point of $L_{\alpha-1}$. Now, since $p(T, p_i) = p(T, T(p_i))$ for all i, it follows that this function is monotone increasing with i. Thus $T(p) \epsilon L_\alpha$ by definition of this set. Therefore, L_α is invariant. It is closed as an immediate result of its definition. It is evident that the initial step in the induction is possible. Consequently, Lemma 1 is established.

LEMMA 2. *For every ordinal α for which L_α is non-vacuous, we have that $L_\alpha - L_{\alpha+1}$ is also non-vacuous.*

The proof is an immediate consequence of a result of Montgomery.[2]

LEMMA 3. *There exists an ordinal α of the first or second number class such that $L_\alpha = 0$. Moreover, the smallest such ordinal is an isolated number.*

The proof is an immediate consequence of Lemma 2 and two well-known theorems for sets with countable bases, namely, the Baire-Hausdorff Theorem and the Durchschnittssatz of Cantor.[3]

2. The principal theorems.

THEOREM I.[4] *If M is a compact metric space and $T(M) = M$ a pointwise periodic transformation, then every closed invariant subset L of M is either vacuous, connected, or has the property that for every separation*

$$L = L^1 + L^2$$

there exists an integer N such that $T^N(L^i) = L^i$ $(i = 1, 2)$.

Proof. If the theorem is not true, there exist a closed invariant subset L^* of M and a separation

$$L = L^1 + L^2 \tag{1}$$

such that for no k is $T^k(L^1) = L^1$. It follows that for every k we have

$$T^k(L^1) \neq L^1 \quad \text{and} \quad T^k(L^2) \neq L^2.$$

Let $L^i_\alpha = L_\alpha L^i$ $(i = 1, 2)$ for every α. Let α be the smallest ordinal such that for some integer k we have $T^k(L^1_\alpha) = L^1_\alpha$. Such an integer k must exist by Lemma 3, and α must be of the first or second number class by the same lemma. Henceforth k shall always denote this integer.

(i) *α is a limit number.*

Otherwise $L_{\alpha-1}$ exists and we consider it to be our space. By Lemma 1, $L_{\alpha-1}$ is closed and compact, hence T^k is uniformly continuous on this set. It

[2] See Deane Montgomery, *Pointwise periodic homeomorphisms*, American Journal of Mathematics, vol. 59(1937), pp. 118-120. See in particular the first sentence of the proof of the lemma of this paper.

[3] See Alexandroff and Hopf, *Topologie*, I, pp. 79 and 85, respectively.

[4] Suggestions by the referee have enabled the authors to shorten materially the proof of this theorem.

follows easily from this that there exists an $\epsilon > 0$ such that[5]

$$T^k(S(L^1_\alpha, \epsilon)) \subset L^1_{\alpha-1} \quad \text{and} \quad T^k(S(L^2_\alpha, \epsilon)) \subset L^2_{\alpha-1}.$$

By definition of L_α we have that $p(T, x)$ is bounded on $L_{\alpha-1} - S(L_\alpha, \epsilon)$; hence let N be its least upper bound on this set.

From the definition of α, we have that $L^1_{\alpha-1}$ is not invariant under $T^{N!k}$. This means that there exists a point x in $L^1_{\alpha-1}$ or $L^2_{\alpha-1}$ such that $T^{N!k}(x)$ is in the other one of these sets. We assume, without loss of generality, that x is in $L^1_{\alpha-1}$. Thus $T^{N!k}(x) \in L^2_{\alpha-1}$. It follows that $p(T, x) > N$, therefore $x \in S(L^1_\alpha, \epsilon)$. Using the fact that every image of x has the same period as x, we have at once that all the points $T^k(x), T^{2k}(x), T^{3k}(x), \cdots, T^{N!k}(x)$ lie in $S(L^1_\alpha, \epsilon)$. This contradicts the definition of x, and completes the proof of (i).

Since α is a limit number of the first or second number class, there must exist an infinite sequence of ordinals $\beta_1 < \beta_2 < \cdots < \alpha$ such that

$$L_\alpha = \prod_{n=1}^{\infty} L_{\beta_n}.$$

From the definition of α we have that for every pair of integers h and n,

$$T^h(L^1_{\beta_n}) \neq L^1_{\beta_n}.$$

If we recall the definition of k, it follows as in the proof of (i) that there exists an $\epsilon > 0$ such that

$$T^k(S(L^1_\alpha, \epsilon)) \subset L^1 \quad \text{and} \quad T^k(S(L^2_\alpha, \epsilon)) \subset L^2.$$

(We are here considering L as the space.) Choose n large enough so that $L^1_{\beta_n}$ is contained in $S(L^1_\alpha, \epsilon)$. Letting x be any point of $L^1_{\beta_n}$, we have that x is in $S(L^1_\alpha, \epsilon)$, hence $T^k(x)$ is in L^1. But $T^k(x)$ is in L_{β_n} since this set is invariant under all powers of T, by Lemma 1. It follows that $T^k(x)$ is in $L^1_{\beta_n}$, whence this set is invariant under T^k. This is impossible as $\beta_n < \alpha$.

This completes the proof of Theorem I.

Notation. By an orbit under T we shall mean the finite set of points consisting of a point x of M and all the images of x under T. A finite set of points X which lie in the same orbit G under T are said to be *consecutive points under* T provided they may be ordered so that $T(x_1) = x_2$, $T(x_2) = x_3$, etc., using the positive integers consecutively.

THEOREM II. *If M is a compact metric space, $T(M) = M$ a pointwise periodic transformation, $\{G_k\}$ a convergent sequence of orbits under T with limit set L, and if there is in L a connected set B such that $T(B) = B$, then L is connected.*

[5] We say that $M = R + S$ is a separation of M provided neither R nor S is vacuous and $\bar{R}S = 0 = R\bar{S}$, where $\bar{N}$ is the closure of a set N. The symbol $S(M, d)$ denotes the set of all points x of a given metric space A containing M such that $\rho(x, m) < d$ for some point $m \in M$.

In order to establish this theorem we shall make use of two lemmas which we now prove. The notation is that of Theorem II.

LEMMA 4. *For any separation of* L

$$L = L^1 + L^2 \quad \textit{with} \quad \rho(L^1, L^2) = 4\epsilon > 0$$

and for any $\delta < \epsilon$ *there exist two integers* r, N *such that for all* $k > N$, $G_kS(L^1, \delta)$ *and* $G_kS(L^2, \delta)$ *are invariant sets under* $S = T^r$.

Proof. Since L is closed, invariant, and disconnected, it follows from Theorem I that there exists an integer r such that L^1 and L^2 are both invariant sets under the transformation $S = T^r$. The existence of the integer N follows at once from the continuity of S and the convergence of the sequence $\{G_k\}$ to L.

LEMMA 5. *If*

$$L = L^1 + L^2 \quad \textit{with} \quad \rho(L^1, L^2) = 4\epsilon > 0$$

is any separation of L, *and* L^1, L^2 *are invariant sets under* $T^r = S$, *then there exists an integer* N' *such that for every* $k > N'$ *the orbit* G_k *under* T *has at most* $(r - 1)$ *consecutive points in* $S(L^i, \delta)$ *for* $i = 1, 2$, *where* δ *is preassigned and less than* ϵ.

Proof. Since $L = \text{Lim}\, G_k$ there exists an integer M such that for all $k > M$, G_k intersects *both* of the sets $S(L^i, \delta)$. Let N' be the larger of the two numbers M and N, where N is chosen as in Lemma 4. Assume that there exists a $k > N'$ such that G_k has r consecutive points p_h $(h = 0, 1, 2, \cdots, r - 1)$ in one of the sets $S(L^i, \delta)$, say in $S(L^1, \delta)$ for definiteness. For any point q in G_k we have $q = T^m(p_0)$ for some integer m, by the definition of G_k. Let $m = nr + t$ define n and t, where these two symbols represent positive integers or zero and $t < r$. Then $q = S^n(p_t)$ by the definition of the transformation S. Consequently, by Lemma 4, we have that $q \in S(L^1, \delta)$, so that, since q was arbitrary, $G_k \subset S(L^1, \delta)$. This contradicts the fact that $k > N' \geq M$ and proves the lemma.

Proof of Theorem II. If the theorem is false, there exists a separation

$$L = L^1 + L^2 \qquad \text{with} \qquad \rho(L^1, L^2) = 4\epsilon > 0.$$

Choose any $\delta < \epsilon$ and the corresponding integers r and N' as given by Lemmas 4 and 5. Let $\{p_i\}$ be a sequence of points converging to $b \in B$ such that for every i we have $p_i \in G_{N'+i} \cdot S(L^1, \delta)$. By Lemma 5, for every i there exists an integer $t_i < r$ such that $T^{t_i}(p_i)$ is not an element of $S(L^1, \delta)$. Since the sequence of integers $\{t_i\}$ assumes only a finite number of different values, infinitely many of the terms of this sequence must be equal to some positive integer $s < r$. Consequently, we may find a subsequence $\{p_i'\}$ of $\{p_i\}$ such that for every i we have $T^s(p_i')$ is not an element of $S(L^1, \delta)$. Since $\lim p_i' = b$ and $T^s(b) = b' \in B \subset S(L^1, \delta)$, this contradicts the fact that the transformation T^s is continuous. This completes the proof of Theorem II.

COROLLARY. *If, in addition to the hypotheses of the theorem, we also assume*

that there exists an integer I such that for every k the orbit G_k contains at most I points, then L consists of exactly one fixed point.

Proof. Let b be any point in B and let x be any point of L which is not in the orbit determined by b. This is possible when the corollary is assumed to be false since both B and L are connected. If we now let $\{b_i\}$ and $\{x_i\}$ denote sequences converging to b and x respectively with the further property that both b_i and x_i belong to the same orbit G_i (for each i), it follows that for every i there exists an integer n_i such that $T^{n_i}(b_i) = x_i$. Since $n_i \leqq I$ for any i, we lose nothing in stating that the sequences have been so selected that $T^s(b_i) = x_i$ for some fixed positive integer s and for every i. But $T^s(b) \epsilon G_b$, the orbit determined by b, while x is not an element of G_b. This is impossible since T^s is continuous and $T^s(b_i) = x_i$ form a sequence converging to x. The proof is thus complete.

Theorem II and its corollary are in their most appealing form (from an intuitive point of view) in the special case in which the set B is a fixed point.

3. **Remarks and example.** Aside from the lemmas and devices used in the proofs above various questions naturally arise. Many of these can be answered in a moment. For example, does connectivity for L imply connectivity for L_1? Conversely, does it follow for L if assumed for L_1? Both of these questions have negative answers. On the other hand, assuming the situation as described in Theorem II and postulating in addition that L contains a fixed point, we ask whether L consists entirely of fixed points and find this question to be more of a challenge. The answer is given in the example below. Similar examples are known where M is a locally connected continuum and L contains points of all finite periods. One further question was easily answered, namely, can L (as in Theorem II) be free of fixed points? Here again the answer is by example, i.e., positive.

Example. In the Euclidean 3-space define an infinite sequence of planes by the equations $P_i : z = i^{-1}$. We shall use cylindrical coördinates (r, θ, z) to define our space M. Let C be the circle $r = 1$, $z = 0$. Define the points $A(1, \pi, 0)$, $B(1, 0, 0)$ and let L consist of C and the segment AB of the polar axis. Thus L is a theta curve.

By definition let $q_k^i = (1 + (1 - k)/i, \pi, 1/i)$ for $k = 1, 2, 3, \cdots, 2i + 1$. Here the range for i is the sequence of positive integers and this is likewise true wherever i occurs below. Furthermore, let $q_{2i+2k}^i = (1, (k\pi)/i, 1/i)$ and $q_{2i+2k+1}^i = (1, (-k\pi)/i, 1/i)$ for $k = 1, 2, 3, \cdots, i - 1$. Now define the orbit G_i by the equation

$$G_i = \sum_{j=1}^{4i-1} q_j^i$$

and subsequently

$$M = L + \sum_i G_i .$$

The transformation T is defined as follows. For every point $p = (r, \theta, 0)$ in L let $T(p) = (r, -\theta, 0)$. For every q_j^i, with $j < 4i - 1$, define $T(q_j^i) = q_{j+1}^i$. Finally, let $T(q_{4i-1}^i) = q_1^i$.

In this example we find that L is a theta curve with one closed free arc consisting of fixed points and all other points in L of period 2 under T.

University of Virginia and University of California at Los Angeles.

COMMUTATIVE ALGEBRAS WHICH ARE POLYNOMIAL ALGEBRAS

By R. F. Rinehart

1. **Introduction.** A polynomial $p(x)$ of non-zero degree with coefficients in an arbitrary field F gives rise to a linear algebra P, with a principal unit, over F. P may be viewed from two standpoints: (1) as the algebra generated by an element x whose minimum equation is $p(x) = 0$; (2) as the algebra of the residue classes modulo $p(x)$ of the ring of all polynomials with coefficients in F. From the first standpoint the elements $1, x, x^2, \cdots, x^{\alpha-1}$, where α is the degree of $p(x)$, constitute a basis for P; from the second standpoint the residue classes $[1], [x], [x^2], \cdots, [x^{\alpha-1}]$ constitute a basis. In either case P, considered as an abstract algebra, has the same properties. We shall call such an algebra *the polynomial algebra generated by* $p(x)$.

This paper had its origin in the speculation as to whether every commutative algebra[1] with a principal unit might be equivalent to a polynomial algebra.[2] The question is here answered in the negative, but it is found that the algebras which are thus completely characterized by polynomials constitute a wide class of commutative algebras. Under proper restrictions, depending on the nature of the ground field, it is shown that the equivalence of a commutative algebra with a principal unit to a polynomial algebra depends only on the structure of the radical. This structure is most conveniently described in terms of the écarts of certain nilpotent subalgebras, a concept which plays a prominent rôle in some recent researches of Scorza.

The results to follow are established under very loose hypotheses on the ground field F, namely, that F is separable, and that F, in case it is finite, has more elements than the commutative algebra in question has indecomposable Peirce components of a common order, for every order. When F is inseparable, the present analysis of the structure of the irreducible polynomial algebra, on which the treatment in §§3 and 4 is fundamentally based, fails. Moreover, as the examples of §5 show, the results of §§3 and 4 are not true for an inseparable ground field without imposition of further restrictions. The writer hopes to investigate the inseparable case later.

2. **Converses of the decomposition theorem for polynomial algebras.** Let P be the algebra generated by the polynomial $p(x)$ with coefficients in F, and let

Received April 20, 1938; presented to the American Mathematical Society, December 29, 1937.

[1] Throughout this paper the term algebra will be used to signify a linear associative algebra of finite order.

[2] Two algebras are said to be equivalent if a simple ring isomorphism exists between the elements of one algebra and the elements of the other.

$p(x) = \prod_{i=1}^{\nu} p_i^{h_i}(x)$, where the $p_i(x)$ are the distinct irreducible factors of $p(x)$ relative to F. It is well known that P is equivalent to the direct sum of the indecomposable polynomial algebras P_i, generated respectively by $p_i^{h_i}(x)$.[3] For the subsequent development a converse of this result is necessary. To that end we prove two lemmas.

LEMMA 1. *A direct sum of indecomposable polynomial algebras over an infinite field is equivalent to a polynomial algebra.*

Proof. Let the polynomial algebras $P_1, P_2, \cdots, P_\nu$ over an infinite field F be generated respectively by the polynomials $p_1^{h_1}(x), p_2^{h_2}(x), \cdots, p_\nu^{h_\nu}(x)$ with coefficients in F, where $p_1(x), p_2(x), \cdots, p_\nu(x)$ are irreducible over F and have unity as their leading coefficients. If the $p_i(x)$ are distinct, then the direct sum

$$P' = P_1 \dotplus P_2 \dotplus \cdots \dotplus P_\nu$$

is equivalent to the algebra P generated by $p(x) = \prod_{i=1}^{\nu} p_i^{h_i}(x)$, since P is equivalent to a direct sum of algebras each of which is equivalent to one and only one of the P_i.

If the $p_i(x)$ are not all distinct, let them be segregated into sets $S_1, S_2, \cdots, S_l$, where each S_k is composed of all the $p_i(x)$ which are transformable into one polynomial of that set by respective transformations of the form $x = y/f_i + g_i$, where $f_i \neq 0$ and g_i are in F, with subsequent multiplication by a power of f_i to produce leading coefficient unity.[4] All of the polynomials of a given set have the same degree. The effect of such transformations on the corresponding algebras P_i is merely one of a change of basis. Consider the set S_r. Let the $p_i(x)$ be so ordered that S_r consists of $p_1(x), p_2(x), \cdots, p_c(x)$. Let the $p_i(x)$ of S_r be carried into $p_i'(y) = p_1'(y) = p_1(y)$ by transformations of the type mentioned. We wish to show that by further transformation the members of S_r can be carried into distinct polynomials.

We note first that if $p_1(y) = y$, the further transformations $y = z + g_i'$ $(i = 1, 2, \cdots, c)$, where the g_i' are distinct elements of F, carry the polynomials of S_r into distinct polynomials. Assuming now that $p_1(y) \neq y$, we distinguish two cases:

Case I. *F is of characteristic zero.*

Case II. *F is of characteristic $q \neq 0$.*

In Case I, F contains a subfield simply isomorphic to the rational field. Hence

[3] See W. Krull, *Algebraische Erweiterungen kommutativer hyperkomplexer Systeme*, Math. Ann., vol. 97, p. 473. The above result follows from Krull's theorem connecting the factorization of the zero ideal of a commutative algebra and the direct sum decomposition of the algebra.

See also A. A. Albert, *Modern Higher Algebra*, Chicago, 1937, pp. 245–249.

This result suffices for the needs of the present paper, but it may be remarked that the slightly more general theorem, in which $p(x)$ is separated into relatively prime factors, also holds, the component algebras being not necessarily indecomposable in such a case.

[4] It will be tacitly understood in the remainder of §2 that after each such transformation each resulting polynomial is to be divided by its leading coefficient.

the application of the further transformations, $y = z/f_i'$ $(i = 1, 2, \cdots, c)$, the f_i being distinct "positive rational integers", will carry the $p_i'(y)$ respectively into polynomials $p_i''(z)$ which differ in their terms of zero degree, since the $p_i(x)$, and consequently the $p_i''(z)$, being irreducible and non-linear, have non-zero constant terms.

Suppose in Case II that every element of F is algebraic with respect to the prime subfield Σ of F. Σ is finite of order q and hence is separable.[5] Since F is infinite and Σ is separable, F contains elements of arbitrarily high degree with respect to Σ. Let ϵ be an element whose degree relative to Σ exceeds cd, where d is the common degree of the $p_i(x)$ of S_r.

On the other hand if F contains an element which satisfies no algebraic equation with coefficients in Σ, let that element be denoted by ϵ.

Now in either event the secondary transformations

$$y = z/\epsilon^i \qquad (i = 1, \cdots, c)$$

carry the $p_i'(y)$ respectively into polynomials $p_i''(z)$ which differ in their terms of degree zero. For, if a_d denotes the constant term of $p_1'(y)$, $p_j''(z) = p_k''(z)$ implies $a_d\epsilon^{jd} = a_d\epsilon^{kd}$, which, since j and k are less than c, implies that $j = k$.

Thus, in either Case I or Case II, distinct generators $p_1^{h_1}(x), \cdots, p_c^{h_c}(x)$ can be produced by transformations of bases of $P_1, \cdots, P_c$. Since a polynomial of one set cannot be carried into a polynomial of another set, the polynomials $p_1(x), \cdots, p_r(x)$ can be made distinct by making the polynomials within each set distinct. Then the argument of the initial paragraph of the proof can be applied.

LEMMA 2. *Let the polynomial algebras $P_1, P_2, \cdots, P_r$ over a finite field F of characteristic q and order q^e be generated respectively by powers of irreducible (in F) polynomials $p_1^{h_1}(x), p_2^{h_2}(x), \cdots, p_r^{h_r}(x)$. If the number of the polynomials $p_i(x)$ which have a common degree u is less than q^e for every u, then the direct sum of the algebras P_i is equivalent to a polynomial algebra.*

Proof. Let the $p_i(x)$ be segregated into sets $S_1, \cdots, S_l$, where each set consists of all the $p_i(x)$ which are transformable into one polynomial of that set by respective transformations of the form $y = z/f_i + g_i$,[6] where $f_i \neq 0$ and g_i are in F. Each such transformation merely effects a change of basis in the corresponding algebra P_i. Consider a set S_r. For convenience let the $p_i(x)$ be so ordered that S_r consists of $p_1(x), \cdots, p_c(x)$. Let each $p_i(x)$ of S_r be carried into $p_i'(y) = p_1'(y) = p_1(y)$ by a corresponding transformation of the type mentioned.

If S_r consists of polynomials whose common degree is not divisible by q, apply the secondary transformations

$$y = z + b_i \qquad (i = 1, 2, \cdots, c),$$

where the b_i are distinct elements of F, to the corresponding $p_i'(y)$. The b_i can be chosen distinct, since, by hypothesis, $c < q^e$. Under these transforma-

[5] See B. L. van der Waerden's *Moderne Algebra*, Berlin, 1930, vol. I, p. 118.

[6] Recall the footnote to the second paragraph of the proof of Lemma 1.

tions $p_1'(y), \cdots, p_c'(y)$ are carried into corresponding polynomials $p_1''(z), \cdots, p_c''(z)$ which are distinct. For if $p_h''(z) = p_k''(z)$,

$$(z + b_h)^n + a_1(z + b_h)^{n-1} + \cdots + a_n = (z + b_k)^n + a_1(z + b_k)^{n-1} + \cdots + a_n ,$$

equating the coefficients of z^{n-1} yields

$$nb_h + a_1 = nb_k + a_1 .$$

Since $n \not\equiv 0 \pmod q$, this implies $b_h = b_k$, which is valid only if $h = k$.

On the other hand, suppose that the polynomials of the set S_r have a common degree divisible by q, say qw. We wish to show that $p_1(y)$ can be carried into at least $q^e - 1$ distinct polynomials by transformations $y = z/f + g$, where $f \neq 0$ and g are in F. We shall assume that this is impossible and show that this assumption leads to a contradiction.

Since $p_1(y)$ is irreducible and since F is separable, the roots of $p_1(y) = 0$ are distinct. Let ϵ be a root of $p_1(y) = 0$. By assumption there is a number $g \neq 0$ of F for which $p_1(z + g) \equiv p_1(z)$. Hence $\epsilon + g$ is also a root of $p_1(y) = 0$. It follows immediately that

$$\epsilon + tg \qquad (t = 0, 1, \cdots, q - 1)$$

are all roots of $p_1(y) = 0$. Again, by assumption, there is a number f of F, different from 0 and 1, for which $f\epsilon$ is a root of $p_1(y) = 0$. It is apparent that $f^i\epsilon$ is also a root for every i. Since $f^{q^e-1} = 1$, and since $f^{q^e-1-r}\epsilon + tg$ is a root, then $\epsilon + tgf^r$ is also a root for every t and r. Finally, from the isomorphism $\epsilon \rightarrow \epsilon^{q^s}$, it follows that ϵ^{q^s} is a root for every s.

From all the elements $f \neq 1$ of F, such that $f\epsilon$ is a root of $p_1(y) = 0$, let b be one for which the smallest positive integer δ, for which $b^\delta = 1$, is minimal. δ divides $q^e - 1$. We consider the following possibilities:

Case I. $\delta < q$;

Case II. $\delta > q$.

If $\delta < q$, we separate the following classes of roots:

$$K_\sigma = \{\epsilon^{q^{e\sigma}} + tg\} \qquad (t = 0, 1, \cdots, q - 1; \sigma = 0, 1, \cdots, w - 1).$$

If $\delta > q$, we make the separation into classes:

$$\bar{K}_\sigma = \{\epsilon^{q^{e\sigma}} + tgb^r\} \quad (t = 0, \cdots, q - 1;\ r = 0, \cdots, \delta - 1;\ \sigma = 0, \cdots, \mu),$$

where μ satisfies the inequalities

$$\mu[(q - 1)\delta + 1] < qw, \qquad (\mu + 1)[(q - 1)\delta + 1] \geqq qw.$$

The number of elements in any K_σ is q, and the number in any $\bar{K}_\sigma$ is $(q - 1)\delta + 1$. We shall show that all the elements of the sets K_σ ($\bar{K}_\sigma$) are distinct.

In any K_σ the roots are obviously distinct. In Case II suppose that two roots of one set were equal. Then we would have

$$t_1b^{r_1} = t_2b^{r_2}. \tag{2.1}$$

If $t_1 = 0$, then $t_2 = 0$, or if $r_1 = r_2$, then $t_1 = t_2$, and in either event the two elements would be identical in $\bar{K}_\sigma$. Hence $r_1 \neq r_2$, say $r_1 > r_2$. Then (2.1) would imply

$$b^{r_1 - r_2} = t_2/t_1,$$

that is, $b^{r_1-r_2}$ would be a number of the prime subfield of F. Since $r_1 - r_2 < \delta$, $b^{r_1-r_2} \neq 1$, and therefore $b^{r_1-r_2}$ would serve as an f of index δ_1 less than q, contrary to the assumption in Case II. Thus the q elements of any K_σ [$(q-1)\delta + 1$ elements of any $\bar{K}_\sigma$] are distinct.

Now suppose that an element of K_h ($\bar{K}_h$) coincided with an element of K_k ($\bar{K}_k$), $h > k$. We shall employ a method indicated in a paper by Ore[7] to show that this assumption implies that $p_1(y)$ is reducible. In either Case I or Case II we would have

$$\epsilon^{q^{eh}} = \epsilon^{q^{ek}} + d, \tag{2.2}$$

where d is in F. From (2.2) one obtains

$$\epsilon^{q^{e(h+1)}} = \epsilon^{q^{e(k+1)}} + d, \tag{2.3}$$

and the elimination of d yields

$$\epsilon^{q^{e(h+1)}} - \epsilon^{q^{eh}} = \epsilon^{q^{e(k+1)}} - \epsilon^{q^{ek}}. \tag{2.4}$$

From (2.4) we form the successive $h - k - 1$ relations,

$$\begin{cases} \epsilon^{q^{e(h+2)}} - \epsilon^{q^{e(h+1)}} = \epsilon^{q^{e(k+2)}} - \epsilon^{q^{e(k+1)}}, \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \epsilon^{q^{e(2h-k)}} - \epsilon^{q^{e(2h-k-1)}} = \epsilon^{q^{eh}} - \epsilon^{q^{e(h-1)}}. \end{cases} \tag{2.5}$$

Addition of the relations (2.4) and (2.5) gives

$$\epsilon^{q^{e(2h-k)}} = 2\epsilon^{q^{eh}} - \epsilon^{q^{ek}}. \tag{2.6}$$

From (2.6) one obtains

$$\epsilon^{q^{e(3h-2k)}} = 2^2\epsilon^{q^{eh}} - 2\epsilon^{q^{ek}} - \epsilon^{q^{eh}} = 3\epsilon^{q^{eh}} - 2\epsilon^{q^{ek}}.$$

It is easily demonstrated by mathematical induction that

$$\epsilon^{q^{e(jh-[j-1]k)}} = j\epsilon^{q^{eh}} - (j-1)\epsilon^{q^{ek}} \qquad (j = 1, 2, \cdots). \tag{2.7}$$

For $j = q$ this yields

$$\epsilon^{q^{e[qh-(q-1)k]}} = \epsilon^{q^{ek}}.$$

Hence $\epsilon^{q^{ek}}$ would satisfy the equation

$$y^{q^{eq(h-k)}} - y = 0, \tag{2.8}$$

and since $\epsilon^{q^{ek}}$ is a root of $p_1(y) = 0$, $p_1(y)$ would divide the left member of (2.8). But in Case I, $q(h-k) < qw$, since $h - k < w$; and in Case II, $q(h-k) <$

[7] *Contributions to the theory of finite fields*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 243–274.

$[(q-1)\delta+1]\mu < qw$. By known irreducibility criteria,[8] this implies that $p_1(y)$ is reducible. This is a contradiction of hypothesis. Hence the elements of the sets K_σ $(\bar{K}_\sigma)$ comprise all of the roots of $p_1(y) = 0$ without repetition.

Now $b\epsilon$ is a root and therefore must be equal to one of the elements of some K_σ $(\bar{K}_\sigma)$, say K_h $(\bar{K}_h)$. $h \neq 0$ in either Case I or Case II, otherwise it would follow that ϵ was an element of F, and this is impossible, since $p_1(y)$ is irreducible and of degree greater than one. Therefore

$$b\epsilon = \epsilon^{q^{eh}} + d, \tag{2.9}$$

where d is in F. From (2.9) we obtain

$$b\epsilon^{q^e} = \epsilon^{q^{e(h+1)}} + d,$$

whence on eliminating d there results

$$b\epsilon^{q^e} - b\epsilon = \epsilon^{q^{e(h+1)}} - \epsilon^{q^{eh}}. \tag{2.10}$$

From (2.10) we form successively the $h-1$ relations

$$\begin{cases} b\epsilon^{q^{2e}} - b\epsilon^{q^e} & = \epsilon^{q^{e(h+2)}} - \epsilon^{q^{e(h+1)}}, \\ b\epsilon^{q^{3e}} - b\epsilon^{q^{2e}} & = \epsilon^{q^{e(h+3)}} - \epsilon^{q^{e(h+2)}}, \\ \quad\vdots \qquad\quad \vdots & \qquad \vdots \qquad\qquad \vdots \\ b\epsilon^{q^{eh}} - b\epsilon^{q^{e(h-1)}} & = \epsilon^{q^{2eh}} - \epsilon^{q^{e(2h-1)}}. \end{cases} \tag{2.11}$$

From the addition of relations (2.10) and (2.11) we find

$$\epsilon^{q^{2eh}} = (b+1)\epsilon^{q^{eh}} - b\epsilon. \tag{2.12}$$

From (2.12) can be obtained

$$\begin{aligned} \epsilon^{q^{3eh}} &= (b+1)^2\epsilon^{q^{eh}} - b(b+1)\epsilon - b\epsilon^{q^{eh}} \\ &= (b^2+b+1)\epsilon^{q^{eh}} - (b^2+b)\epsilon. \end{aligned}$$

By mathematical induction it is easily shown that

$$\epsilon^{q^{e\delta h}} = (b^{\delta-1} + b^{\delta-2} + \cdots + 1)\epsilon^{q^{eh}} - (b^{\delta-1} + b^{\delta-2} + \cdots + b)\epsilon,$$

whence, since $b^\delta = 1$, and $b \neq 0, 1$,

$$\epsilon^{q^{e\delta h}} = \epsilon.$$

Thus ϵ satisfies the equation

$$y^{q^{e\delta h}} - y = 0, \tag{2.13}$$

and therefore the left member of (2.13) is divisible by $p_1(y)$. However, in Case I, $\delta < q$ and $h < w$, so that $h\delta < qw$; and in Case II, $h\delta \leq \mu\delta < \mu[(q-1)\delta+1] < qw$. But $h\delta < wq$ implies, as before, that $p_1(y)$ is reducible. Consequently our original assumption is untenable, and since the number of the $p_i'(y)$ of S_r

[8] See O. Ore, loc. cit.

is less than $q^e - 1$, the $p'_i(y)$ can be carried into distinct polynomials $p''_i(z)$ by transformations of admissible type.

Since a polynomial of one set cannot be carried into a polynomial of another set by transformations of the type employed, all the $p_i(x)$ can be made distinct by making those of the individual sets distinct. Then the argument of the first paragraph of the proof of Lemma 1 can be used, and the proof of Lemma 2 is complete.

It is apparent that two minor refinements, which need not be mentioned here, could be made in the statement of Lemma 2. However, examples can be easily constructed to show that the lemma, as stated, is false if the number of the $p_i(x)$ having a common degree exceeds $q^e - 1$, so that the condition stated is the best possible of its type.

3. **Commutative algebras equivalent to polynomial algebras.** The importance of the preceding lemmas for the purpose of this paper lies in the fact that criteria for the equivalence of an indecomposable commutative algebra to a polynomial algebra will thereby furnish criteria for the equivalence of a decomposable commutative algebra to a polynomial algebra. Let C be an arbitrary commutative algebra, with a principal unit, over a field F. By the classical Peirce decomposition C is representable as a direct sum of commutative algebras C_i $(i = 1, 2, \cdots, \nu)$, where each C_i has a principal unit but no other idempotent.[9] By a well known theorem of linear algebras,[10] if the field F is separable, such an algebra C_i is the sum of a field D_i and the radical N_i of C_i. We shall henceforth assume that F is a separable field. Since D_i has a finite basis, $d_1, d_2, \cdots, d_\beta$, over F, and since F is separable, there is a polynomial equation $c_i(\lambda) = 0$ of degree β, which is irreducible in F and which is satisfied by some element x of D_i.[11] Hence $1, x, x^2, \cdots, x^{\beta-1}$ constitute a basis for D_i.

The radical N_i of C_i can be considered as an algebra $\bar{N}_i$ over D_i.[12] For if $n_1, n_2, \cdots, n_s$ is any maximal set of elements of N_i which are linearly independent with respect to D_i, then the elements

$$x^j n_k \qquad (j = 0, 1, \cdots, \beta - 1; k = 1, 2, \cdots, s) \tag{3.1}$$

are linearly independent over F. Since any element of N_i which was linearly independent over F of the set (3.1) would be independent of $n_1, \cdots, n_s$ over D_i, the set (3.1) constitutes a basis for N_i. Thus the algebra with the basis $n_1, n_2, \cdots, n_s$ over D_i contains every element of N_i, and conversely. We note as a special case that the radical of the indecomposable polynomial algebra

[9] See, for example, G. Scorza, *Sulle algebre riducibili*, Rend. del Sem. Math. d. R. Univ. di Roma, Ser. IV, vol. 1(1937), pp. 188–189.

[10] M. Deuring, *Algebren*, Berlin, 1935, p. 23; and L. E. Dickson, *Algebren und ihre Zahlentheorie*, Zurich, 1927, p. 32.

[11] B. L. van der Waerden, op. cit., pp. 118–120.

[12] If C_i has no radical, it is obviously a polynomial algebra.

generated by $p^k(x)$, where $p(x)$ is irreducible in F, has a basis of the form

$$x^j p^k(x) \qquad (j = 0, 1, \cdots, \beta - 1; k = 1, 2, \cdots, h);$$

that is, the n_k can be chosen as powers of a single element of the radical.

We are now in a position to characterize the indecomposable polynomial algebra by means of the écart of its radical. The écart of an algebra A is defined to be the difference between the orders of A and A^2. Scorza has shown[13] that if α is the index,[14] r the order, and δ the écart of a commutative nilpotent algebra, then

$$\binom{\alpha + \delta - 1}{\delta} - 1 \geqq r. \tag{3.2}$$

Now, in the indecomposable algebra C_i, $(\bar{N}_i)^2$ is the same algebra as $\overline{(N_i^2)}$, since each algebra is composed of all linear combinations over F of elements of the form $(d_p n_p)(d_q n_q)$, where d_p and d_q are in D_i. Let δ be the écart of N_i and $\bar{\delta}$ the écart of $\bar{N}_i$. Then, as in the case of N_i and $\bar{N}_i$, if $n_1', n_2', \cdots, n_{s-\bar{\delta}}'$ form a basis for $\bar{N}_i^2$, the elements

$$x^j n_k' \qquad (j = 0, 1, \cdots, \beta - 1; k = 1, 2, \cdots, s - \bar{\delta})$$

form a basis for N_i^2. Hence

$$\delta = \beta\bar{\delta}. \tag{3.3}$$

In case C_i is a polynomial algebra, then the écart of N_i is equal to the order of D_i. This condition is also sufficient to insure that an indecomposable commutative algebra be a polynomial algebra; for the equality of δ and β implies, from (3.3), that $\bar{\delta} = 1$, and Scorza's inequality (3.2) becomes

$$\binom{\alpha}{1} - 1 = \alpha - 1 \geqq r,$$

where α is the index of $\bar{N}_i$. But $\alpha - 1$ cannot exceed r,[15] hence $\alpha - 1 = r$. Frobenius has shown that if the index of a nilpotent algebra exceeds its order by unity, then the algebra has a basis of the form $n, n^2, \cdots, n^{\alpha-1}$. Hence C_i is of the form $D_i + N_i$, where D_i is generated by an irreducible polynomial $c_i(x)$ and N_i has a basis of the form

$$x^j n^k \qquad (j = 0, 1, \cdots, \beta - 1; k = 1, 2, \cdots, \alpha - 1).$$

C_i is therefore equivalent to the polynomial algebra generated by the polynomial $c_i^\alpha(x)$. Hence from Lemmas 1 and 2 we can state

THEOREM 3.1. *Let C be a commutative algebra with a principal unit over a separable field F, and let $C_1, \cdots, C_v$ be the indecomposable components in a Peirce*

[13] G. Scorza, *Sulla struttura delle algebre pseudonulle*, Atti Accad. naz. Lincei Rend., Ser. 20, vol. 6(1934), p. 143.

[14] That is, α is the smallest positive integer for which $A^\alpha = 0$.

[15] See, for example, L. E. Dickson, op. cit., pp. 110-111.

decomposition of C. If the number of elements of F exceeds the number of field components D_i (or difference algebras C_i/N_i) which have a common order, then a necessary and sufficient condition that C be equivalent to a polynomial algebra is that the radical N_i of each C_i either be zero or have an écart equal to the order of D_i (or C_i/N_i).

Since C is commutative and has a principal unit, the number of the C_i is equal to the number of primitive idempotents of C, and hence we have

COROLLARY 3.11. *Let C be a commutative algebra with a principal unit over a separable field F. If F has more elements than C has primitive idempotents, the necessary and sufficient condition of Theorem 3.1 holds.*

COROLLARY 3.12. *A semisimple commutative algebra C over a separable field F, which has more elements than C has simple invariant subalgebras of any common order, is equivalent to a polynomial algebra.*

If F is a finite field, then it is separable. Let A be an algebra with a principal unit over F which is such that every primitive idempotent of A is commutative with the radical of A. If A possesses no total matric subalgebra, then each component A_i of a direct sum decomposition of A can be represented as the sum of a division algebra and the radical of A_i.[16] As is well known, every division algebra over a finite field is commutative.[17] Now if the écart of the radical of A_i is equal to the order of the field component of A_i, then the radical of A_i is commutative, as can be seen from the proof of Theorem 3.1. Hence we have

THEOREM 3.2. *Let A be an algebra of the type described above with a principal unit over a finite field F, and let $A_1, \cdots, A_r$ be the indecomposable components in a direct sum decomposition of A. Let N_i be the radical of A_i, and let the number of elements of F exceed the number of the difference algebras A_i/N_i which have a common order. If each N_i is zero or has an écart equal to the order of A_i/N_i, then A is equivalent to a polynomial algebra and is therefore commutative.*

This theorem is an extension of Wedderburn's theorem,[17] which could now be viewed as the special case of Theorem 3.2 in which A is simple. The formulation of corollaries analogous to Corollaries 3.11 and 3.12 is obvious and need not be made here.

Since the écart of a direct sum of algebras is clearly equal to the sum of the écarts of the component algebras, we can state from our analysis of the polynomial algebra

THEOREM 3.3. *Let C be a commutative algebra with a principal unit over a field K, and let C have the radical N. A necessary condition that C be equivalent to a polynomial algebra is that the écart of N shall not exceed the order of the algebra C/N.*

If F is a field of the type mentioned in Theorem 3.1, and if every indecomposable component in the Peirce decomposition of C possesses a radical, then

[16] L. E. Dickson, op. cit., p. 130.

[17] J. H. M. Wedderburn, Trans. Amer. Math. Soc., vol. 6(1905), pp. 349-352.

the equality of the écart of the radical N of C and the order of C/N is a sufficient condition that C be equivalent to a polynomial algebra. Hence we have

THEOREM 3.4. *Let C be a commutative algebra with a principal unit over a separable field F subject to the restriction of Theorem* 3.1, *and let C have the radical N. If C possesses no non-nilpotent simple invariant subalgebra, and if the order of C/N is equal to the écart of N, then C is equivalent to a polynomial algebra.*

The analogue of Theorem 3.2 is evident and will not be stated here.

Remark 1. Theorem 3.4 and Corollary 3.12 give sufficiency conditions that two very diverse types of algebras be equivalent to polynomial algebras. Indeed, it is evident that a large class of commutative algebras comes under the head of polynomial algebras, and with very weak restrictions on the ground field.

Remark 2. Theorem 3.1 may be viewed in the following light. Suppose that there is prescribed a structure type $C_1 \dotplus C_2 \dotplus \cdots \dotplus C_r$, where $\dotplus$ denotes direct sum and each C_i is indecomposable, in which certain of the C_i's have no radicals and the remaining C_i's do possess radicals. Consider the former set of C_i's as playing the rôle of constants, and the latter as playing the rôle of parameters. Let the orders of the C_i/N_i be fixed for the "parameters". Then the écart of the radical of any algebra having this structure type is a function of the "parameters"; and of the totality of commutative algebras with a principal unit, and fitting this structure, the ones which are equivalent to polynomial algebras are those for which the "parameters" minimize the écart of the radical of the algebra.

4. Criteria involving the discriminant matrix.

If a further restriction is placed on the ground field F, the conditions of the theorems of the preceding section can be replaced by more convenient criteria involving the discriminant matrix of C. The discriminant matrix of a commutative algebra, relative to a given basis, is unique. Let us recall the following two properties of it:

(a) If A is an algebra of order n over a field F, and if F is of characteristic 0 or q, $q > n$, then the nullity of the discriminant matrix of A is equal to the order of the radical of A and is independent of the choice of basis of A.[18]

(b) For a proper choice of basis (namely, for a basis consisting of a set of bases each of which is a basis for a component) the discriminant matrix is a direct sum of the discriminant matrices of the components.

From Theorem 3.1 and property (a) follows

THEOREM 4.1. *Let C be a commutative algebra with a principal unit over a field F which is subject to the restrictions of Theorem* 3.1. *Further, if F is finite, let the characteristic of F exceed the order of every indecomposable component C_i. Then a necessary and sufficient condition that C be equivalent to a polynomial algebra is that the radical of each C_i either be zero or have an écart equal to the rank of the discriminant matrix of C_i.*

[18] L. E. Dickson, op. cit., pp. 109–110. The theorem and proof here found are invalid if F is of characteristic q, $0 < q \leqq r$, where r is the rank of the algebra.

C. C. MacDuffee, Annals of Mathematics, (2), vol. 32(1932), pp. 60–66.

Theorem 4.2, the corresponding analogue of Theorem 3.2, is readily formulated and will be omitted here. From properties (a) and (b) and Theorem 3.3 follows

THEOREM 4.3. *Let C and F be as prescribed in Theorem 4.1. A necessary condition that C be equivalent to a polynomial algebra is that the écart of the radical of C shall not exceed the rank of the discriminant matrix of C.*

Properties (a) and (b) and Theorem 3.4 yield

THEOREM 4.4. *Let C be a commutative algebra of order n with a principal unit over a separable field F whose characteristic is 0 or q, $q > n$. If C has no non-nilpotent simple invariant subalgebra, and if the écart of the radical of C is equal to the rank of the discriminant matrix of C, then C is equivalent to a polynomial algebra.*

Similarly, the unstated Theorem 3.5 becomes

THEOREM 4.5. *Let A be an algebra of the type described in Theorem 3.2 of order n with a principal unit over a finite field F of characteristic $q > n$. Let A possess no total matric subalgebra and no non-nilpotent simple invariant subalgebra. If the écart of the radical of A is equal to the rank of the discriminant matrix of A then A is equivalent to a polynomial algebra and is therefore commutative.*

5. **Conclusion**. The following two examples show that the condition that F be separable, employed throughout the investigation, is an essential restriction; that is, it is a restriction inherent in the problem and not peculiar to the method of proof.

Example 1. Let F be the field formed by the adjunction of two independent indeterminates, α and β, to the modular field with modulus 2. This field is inseparable. Let the fields A and B over F be defined by the equations $x^2 - \alpha = 0$ and $y^2 - \beta = 0$, respectively. The direct product, C, of A and B is a commutative algebra with a principal unit, and $1, x, y, xy$ constitute a basis for C. Let

$$\xi = f_1 + f_2x + f_3y + f_4xy,$$

where the f_i are in F, be any element of C. The first matrix of ξ is

$$R(\xi) = \begin{Vmatrix} f_1 & f_2\alpha & f_3\beta & f_4\alpha\beta \\ f_2 & f_1 & f_4\beta & f_3\beta \\ f_3 & f_4\alpha & f_1 & f_2\alpha \\ f_4 & f_3 & f_2 & f_1 \end{Vmatrix},$$

whence the characteristic equation of ξ is found to be

$$\lambda^4 + f_1^4 + f_2^4\alpha^2 + f_3^4\beta^2 + f_4^4\alpha^2\beta^2 = 0. \tag{5.1}$$

The constant term of (5.1) is the square of

$$\zeta = f_1^2 + f_2^2\alpha + f_3^2\beta + f_4^2\alpha\beta.$$

Assume that $\zeta = 0$. Let $f_i = g_i/h_i$, $h_i \neq 0$, $(i = 1, 2, 3, 4)$, where g_i and h_i are polynomials in α and β. Then $\zeta = 0$ implies

$$g_1^2h_2^2h_3^2h_4^2 + g_2^2h_1^2h_3^2h_4^2\alpha + g_3^2h_1^2h_2^2h_4^2\beta + g_4^2h_1^2h_2^2h_3^2\alpha\beta = 0. \tag{5.2}$$

Now if the sum of the first and second terms and the sum of the third and fourth terms of (5.2) possessed degrees, then the first sum would be of even degree in β and the second sum would be of odd degree in β. On the other hand, if one of these sums is zero, the other is also zero. Hence in either case the implication is, since $h_1h_2h_3h_4 \neq 0$,

$$g_1^2h_2^2 + g_2^2h_1^2\alpha = 0, \qquad g_3^2h_4^2 + g_4^2h_3^2\alpha = 0. \tag{5.3}$$

Applying the foregoing argument on each of the equalities (6) with α in place of β, we find that $g_1 = g_2 = g_3 = g_4 = 0$, whence $f_1 = f_2 = f_3 = f_4 = 0$. Thus ζ can be zero only if ξ is zero. Therefore ξ has an inverse, if $\xi \neq 0$, and C is a field and is therefore simple. If Theorem 3.1 held, C would therefore be equivalent to a polynomial algebra. However, the left member of (5.1) is reducible in F, being the square of $\lambda^2 + \zeta$, so that every element of C satisfies an equation of degree 2. Therefore C, being of order 4, cannot be equivalent to a polynomial algebra.

Example 2. Let Q be the algebra defined by $Q = C + N$, where a basis for N is n, xn, yn, xyn, where $n^2 = 0$, and C is the algebra of Example 1. It is easily seen that Q has no simple non-nilpotent invariant subalgebra, and that the écart of N is 4. If Theorems 3.1 or 3.4 held, Q would be equivalent to a polynomial algebra. However, any element, $q = \xi + p$, of Q, where ξ and p are elements of C and N, respectively, satisfies the equation

$$\lambda^2(\lambda^2 + \zeta) = 0$$

of degree 4. Therefore Q, being of order 8, cannot be equivalent to a polynomial algebra.

Remark 3. The theorems of this paper can be rephrased to give conditions under which the rank of an algebra is equal to its order, since the rank is equal to the order, if and only if the algebra is equivalent to a polynomial algebra.

Remark 4. These results can be coupled with a paper by L. Okunew (*Ring als Algebra über einem Korper*, Rec. Math. Moscou, vol. 40, pp. 410–423) to furnish criteria that a general ring be equivalent to a residue class ring modulo a polynomial in one indeterminate with coefficients in a field.

CASE SCHOOL OF APPLIED SCIENCE.

PRÜFER IDEALS IN COMMUTATIVE RINGS

By D. M. Dribin

1. **Introduction.** H. Prüfer has given[1] a general definition of an ideal in a field and has investigated the properties of these ideals in certain ideal systems. In the present paper a similar study is made, but the algebraic domain of reference will be taken to be a commutative ring $\mathfrak{R}$ having a unit element and possessing no divisors of zero.[2]

2. **Divisibility properties of elements.** The present section, although of some interest, is largely irrelevant to the main matter of the paper but can be conveniently treated at this point.

Let $\mathfrak{g}$ be a subring of $\mathfrak{R}$ with a unit element; the concept of divisibility can now be defined *relative to* $\mathfrak{g}$ so that the elements of $\mathfrak{g}$ may be thought of as the *integral elements* of $\mathfrak{R}$. If $a \neq 0$ and $b \neq 0$ are elements of $\mathfrak{g}$, then *a is divisible by b* if $a = bc$, where c is in $\mathfrak{g}$. Obviously, divisibility relative to $\mathfrak{g}$ is a reflexive and transitive property. If a and b divide each other, $a = b\epsilon_1$, $b = a\epsilon_2$, then $\epsilon_1\epsilon_2 = 1$, where ϵ_1 and ϵ_2 are integral elements; such integers which are divisors of 1 are called *units in* $\mathfrak{g}$ and elements a and b related as above, *associated elements*.

If a and b are integral, then an element d in $\mathfrak{g}$ is said to be a *greatest common divisor of a and b* if a and b are divisible by d and if d is divisible by every common divisor of a and b. If d is a unit, then a and b are said to be *relatively prime*. $\mathfrak{R}$ is *complete*[3] (relative to $\mathfrak{g}$) if every pair of elements in $\mathfrak{g}$ has a g. c. d.

A *prime element* p in $\mathfrak{g}$ is an integral element that is not a unit and whose divisors are associated with 1 or p. $\mathfrak{R}$ is *primary* (relative to $\mathfrak{g}$) if for every two integers a and b it is true that either a and b are relatively prime or that there exists a common prime element divisor p of a and b. Hence, if $\mathfrak{R}$ is primary, every integer $a \neq 0$ is either a unit or is divisible by a prime element.

The following theorem is proved in a manner very similar to that of a theorem of Prüfer:[4]

Theorem 1. *If $\mathfrak{R}$ is complete relative to $\mathfrak{g}$, and if $a = a_1 \cdots a_n$ (where a, a_i $(i = 1, \cdots, n)$ are integers) is divisible by b, then $b = b_1 \cdots b_n$, where b_i $(i = 1, \cdots, n)$ is an integer which divides a_i.*

Received May 5, 1938; presented to the American Mathematical Society, April 16, 1938. The author is a National Research Fellow.

[1] *Untersuchungen über Teilbarkeitseigenschaften in Körpern*, Journal für Mathematik, vol. 168(1932), pp. 1–36.

[2] That is, $\mathfrak{R}$ is a domain of integrity (Integritätsbereich) with unit element.

[3] Prüfer, op. cit., p. 3.

[4] Loc. cit., Theorem 3.

Corollary. *If $\mathfrak{R}$ is complete relative to $\mathfrak{g}$, then a prime element p divides a product of two integers if and only if it divides one of the factors.*

Theorem 2. *If $\mathfrak{R}$ is primary relative to $\mathfrak{g}$ and if the number of units in $\mathfrak{g}$ is finite, then there are infinitely many prime elements in $\mathfrak{g}$.*

For, if $p_1, p_2, \cdots, p_n$ were prime elements and if we wrote $\pi = p_1p_2 \cdots p_n$, then not all the elements $\pi^i + 1$ $(i = 1, 2, \cdots)$ are units in $\mathfrak{g}$. For, then we would have $\pi^i + 1 = \pi^j + 1$ for some values of i and j, contrary to the fact that π is a product of prime elements. Hence, for some t, $\pi^t + 1$ is not a unit and is therefore divisible by a prime element p'. Since $p_1, p_2, \cdots, p_n$ do not divide $\pi^t + 1$, it follows that $p' \neq p_i$ $(i = 1, 2, \cdots, n)$.

By the corollary to Theorem 1 we easily prove

Theorem 3. *The set (not necessarily finite) of prime elements that divide an integral element of a complete and primary ring is unique.*

Theorem 4. *If an integral element $a \neq 0$ is divisible by a finite number of distinct prime elements, then each of these prime elements occurs in a to a finite power.*

For, otherwise we could write $a = a_1a_2$, where a_1 contains all the prime elements p which divide a to an arbitrarily high power, and where a_2 is prime to such elements p. But then, as is easily seen, $a_1 = a_1^2$, whence either $a_1 = 0$ or $a_1 = 1$. But neither possibility can occur, since $a \neq 0$ and a_1 is divisible by at least one prime element.

3. **Ideals and ideal systems in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$.** In the following, $\mathfrak{o}$ is a subring of $\mathfrak{g}$ with the unit element. We define an *ideal in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$* (briefly, *ideal*) as follows:[5]

An ideal $\mathfrak{a} = (a_1, a_2, \cdots, a_n)$ in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$ is any set of elements of $\mathfrak{g}$ satisfying the following properties:

(1) *the elements $a_1, a_2, \cdots, a_n$ are contained in $\mathfrak{a}$,*

(2) *if a and b are in $\mathfrak{a}$, so also is $a + b$,*

(3) *(a) is defined to be the set of all elements ar, where r is an element of $\mathfrak{o}$ (principal ideal),*

(4) *$(a_1, \cdots, a_m) \subseteq (b_1, \cdots, b_n)$ if $a_1, \cdots, a_m$ are in $(b_1, \cdots, b_n)$,*

(5) *if a is in $(a_1, \cdots, a_n)$, then ab is in $(a_1b, \cdots, a_nb)$.*

This definition of ideal does not, of course, give a unique meaning to $(a_1, \cdots, a_n)$ but provides, instead, *ideal systems* in each of which $(a_1, \cdots, a_n)$ is defined in a prescribed manner. Later on we shall discuss three different types of ideal systems $\mathfrak{L}$, $\mathfrak{A}$, $\mathfrak{U}$; we denote by $\mathfrak{M}$ the generic ideal system and understand in what follows that we are concerned with such a fixed ideal

[5] It may have been remarked by the reader that the elements of $\mathfrak{g}$ are the only elements of $\mathfrak{R}$ that play any rôle in this theory. That is quite true, and for all the difference it would make, the larger ring $\mathfrak{R}$ could be omitted from consideration; that is, however, a matter of taste and I have chosen not to do so. For, $\mathfrak{g}$ is an arbitrary subring with unit element of $\mathfrak{R}$ and in a possible discussion of the interrelations between two such subrings $\mathfrak{g}$ and $\mathfrak{g}'$ the ring $\mathfrak{R}$ may not be as superfluous as it appears to be here. The theory of (maximal) orders in an algebra is a case in point.

system. We should mention in passing that in cases where ambiguity might arise we write $\mathfrak{a}_{\mathfrak{M}} = (a_1, \cdots, a_n)_{\mathfrak{M}}$ for an ideal in $\mathfrak{M}$.

If $\mathfrak{a} = (a_1, a_2, \cdots) = (a_1', a_2', \cdots)$, $\mathfrak{b} = (b_1, b_2, \cdots) = (b_1', b_2', \cdots)$, then it is easily seen that

$$(a_1b_1, a_2b_1, \cdots, a_1b_2, a_2b_2, \cdots) = (a_1'b_1', a_2'b_1', \cdots, a_1'b_2', a_2'b_2', \cdots),$$

so that we may define the *product* of two ideals in a given ideal system as follows:

If $\mathfrak{a} = (a_1, a_2, \cdots, a_m)$, $\mathfrak{b} = (b_1, b_2, \cdots, b_n)$ are two ideals in an ideal system $\mathfrak{M}$ of $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$, then the *product* $\mathfrak{ab}$ is given by

$$\mathfrak{ab} = (a_1b_1, a_2b_1, \cdots, a_1b_2, a_2b_2, \cdots, a_mb_n).$$

$\mathfrak{ab}$ is an ideal in $\mathfrak{M}$.

The ideal $\mathfrak{a}$ is said to be *divisible by* the ideal $\mathfrak{b}$ if there exists an ideal $\mathfrak{x}$ such that $\mathfrak{a} = \mathfrak{bx}$. If $\mathfrak{a}$ is divisible by $\mathfrak{b}$, then we cannot say, as in the classical theory, that $\mathfrak{a} \subseteq \mathfrak{b}$ but simply that $\mathfrak{a} \subseteq \mathfrak{b}_{\mathfrak{g}}$, where $\mathfrak{b}_{\mathfrak{g}}$ is the linear extension of $\mathfrak{b}$ in $\mathfrak{g}$. That is, if $\mathfrak{b} = (b_1, \cdots, b_n)$, $\mathfrak{b}_{\mathfrak{g}}$ is the set of all $\sum_{i=1}^{n} g_ib_i$, where the g_i are arbitrary elements of $\mathfrak{g}$.

We still have that $\mathfrak{a} \subseteq \mathfrak{b}$ implies $\mathfrak{ab} \subseteq \mathfrak{bc}$ for every ideal $\mathfrak{c}$, whence $a \subseteq \mathfrak{a}$, $b \subseteq \mathfrak{b}$ implies $ab \subseteq \mathfrak{ab}$.

If we define a *prime ideal* $\mathfrak{p}$ to be one having no ideal divisors other than $\mathfrak{p}$ and $\mathfrak{o}$, then we have

THEOREM 5. *In a complete ring $\mathfrak{R}$ the ideal (p) is a prime ideal if p is a prime element.*

Suppose that $(p) = \mathfrak{ab}$, where $\mathfrak{a}$ and $\mathfrak{b}$ are distinct from $\mathfrak{o}$ and (p). Then there exist elements a and b in $\mathfrak{a}$ and $\mathfrak{b}$, respectively, such that neither a nor b is in (p). But ab is in (p) and must be divisible by p, contrary to the corollary of Theorem 1.

By Theorem 2, we have the

COROLLARY. *If the number of units in $\mathfrak{R}$ (relative to $\mathfrak{g}$) is finite, there exist infinitely many prime (principal) ideals in $\mathfrak{R}$ relative to $\mathfrak{g}$ and any subring $\mathfrak{o}$ of $\mathfrak{g}$.*

4. The properties $A, \cdots, E$ of an ideal system $\mathfrak{M}$. The monotonic theorems. If $\mathfrak{M}$ is an ideal system in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$, then we shall say:

$\mathfrak{M}$ has property A if every ideal in $\mathfrak{M}$ is a principal ideal.

$\mathfrak{M}$ has property B if for every $\mathfrak{a}$ and $\mathfrak{b}$ in $\mathfrak{M}$, $\mathfrak{a} \subseteq \mathfrak{b}_{\mathfrak{g}}$ implies $\mathfrak{a} = \mathfrak{bx}$ with $\mathfrak{x}$ in $\mathfrak{M}$.

It is obvious that A implies B.

$\mathfrak{M}$ has property C if for every ideal $\mathfrak{a}$ in $\mathfrak{M}$ there exists an ideal $\mathfrak{b}$ in $\mathfrak{M}$ such that $\mathfrak{ab} = (\gamma)$, where γ is in $\mathfrak{g}$.

B implies C. For, if $\mathfrak{a} = (a_1, \cdots, a_n)$, $(a_1) \subseteq \mathfrak{a}_{\mathfrak{g}}$, whence $(a_1) = \mathfrak{ab}$.

The following condition is equivalent to property C: *for every two ideals $\mathfrak{a}$ and $\mathfrak{b}$ in $\mathfrak{M}$ there exists an element γ in $\mathfrak{g}$ such that $\mathfrak{a}$ divides $(\gamma)\mathfrak{b}$.*

For, let $\mathfrak{ab}' = (\gamma)$; then $\mathfrak{a}(\mathfrak{bb}') = (\gamma)\mathfrak{b}$. Conversely, take $\mathfrak{b} = \mathfrak{o}$. Then there is an integral element γ and an ideal $\mathfrak{b}'$ such that $\mathfrak{ab}' = (\gamma)\mathfrak{b} = (\gamma)\mathfrak{o} = (\gamma)$.

$\mathfrak{M}$ *has property D* if $\mathfrak{a}\mathfrak{b} \subseteq \mathfrak{b}\mathfrak{c}$ ($\mathfrak{c} \neq (0)$) implies $\mathfrak{a} \subseteq \mathfrak{b}$, for every $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ in $\mathfrak{M}$.

C implies D. For, if $\mathfrak{c}\mathfrak{c}' = (\gamma)$, we have $\mathfrak{a}(\gamma) \subseteq \mathfrak{b}(\gamma)$, whence $\mathfrak{a} \subseteq \mathfrak{b}$.

D obtains if $\mathfrak{a}\mathfrak{c} = \mathfrak{b}\mathfrak{c}$, $\mathfrak{c} \neq (0)$, implies that $\mathfrak{a} = \mathfrak{b}$ for every $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ in $\mathfrak{M}$. For, if $\mathfrak{a}\mathfrak{c} \subseteq \mathfrak{b}\mathfrak{c}$, we have $(\mathfrak{a}, \mathfrak{b})\mathfrak{c} = (\mathfrak{a}\mathfrak{c}, \mathfrak{b}\mathfrak{c}) = \mathfrak{b}\mathfrak{c}$, whence $(\mathfrak{a}, \mathfrak{b}) = \mathfrak{b}$, or $\mathfrak{a} \subseteq \mathfrak{b}$.

$\mathfrak{M}$ *has property E* if $(a)\mathfrak{a} \subseteq (b)\mathfrak{a}$, $\mathfrak{a} \neq (0)$ implies that $a = bm$, where m is in $\mathfrak{o}$, for all a and b in $\mathfrak{g}$ and every ideal $\mathfrak{a}$ in $\mathfrak{M}$.

Obviously, D implies E.

THEOREM 6. *Every complete primary ring $\mathfrak{R}$ having the property C for a given ideal system $\mathfrak{M}$ (that is, $\mathfrak{M}C$ holds) has the property $\mathfrak{M}A$, provided that* (1) *if for any ideal $\mathfrak{a}$ it is true that $\mathfrak{a}^2$ divides $\mathfrak{a}$, then $\mathfrak{a} = \mathfrak{a}^2$, and* (2) *every unit in $\mathfrak{g}$ is also in $\mathfrak{o}$.*

Let $\mathfrak{a} = (a_1, \cdots, a_n)$; to prove the theorem it is obviously sufficient to assume that the greatest common divisor of $a_1, \cdots, a_n$ is a unit. If $\mathfrak{a}\mathfrak{b} = (\gamma)$, $\mathfrak{b} = (b_1, \cdots, b_m)$ and if d is the g. c. d. of $b_1, \cdots, b_m$, $b_i = db_i'$, then $\mathfrak{a}\mathfrak{b}'(d) = (\gamma)$, where $\mathfrak{b}' = (b_1', \cdots, b_m')$. Hence $\gamma = \alpha d$, where α is in $\mathfrak{g}$, and we have $\mathfrak{a}\mathfrak{b}'(d) = (\alpha d) = (\alpha)(d)$, whence $\mathfrak{a}\mathfrak{b}' = (\alpha)$. If p were a prime element dividing α, p would divide the $a_i b_j'$ ($i = 1, \cdots, n$; $j = 1, \cdots, m$), whereas we have assumed that g. c. d. of $a_1, \cdots, a_n =$ g. c. d. of $b_1', \cdots, b_m' =$ unit. Hence α is a unit and by hypothesis (2) of the theorem, $(\alpha) = \mathfrak{o} = \mathfrak{a}\mathfrak{b}'$. Hence $\mathfrak{a}^2\mathfrak{b}' = \mathfrak{a}$ and $\mathfrak{a}^2$ divides $\mathfrak{a}$ and by hypothesis (1) $\mathfrak{a}^2 = \mathfrak{a}$. Therefore $\mathfrak{a} = \mathfrak{o}$, since $\mathfrak{M}$ has property D.

Consider two ideal systems $\mathfrak{M}_1$ and $\mathfrak{M}_2$ in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$ such that every ideal $(a_1, \cdots, a_n)$ in $\mathfrak{M}_1$ is contained in the corresponding ideal in $\mathfrak{M}_2$—we write $(a_1, \cdots, a_n)_1 \subseteq (a_1, \cdots, a_n)_2$. We can prove the following monotonic divisibility property:

THEOREM 7. *Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two ideal systems in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$ such that $(a_1, \cdots, a_n)_1 \subseteq (a_1, \cdots, a_n)_2$ for all ideals $(a_1, \cdots, a_n)_1$, $(a_1, \cdots, a_n)_2$. Then an ideal relation $\mathfrak{a}_1 = \mathfrak{b}_1\mathfrak{x}_1$ implies $\mathfrak{a}_2 = \mathfrak{b}_2\mathfrak{x}_2$.*

Write $\mathfrak{a}_1 = (a_1, \cdots, a_t)_1 = \mathfrak{b}_1\mathfrak{x}_1 = (\cdots, b_i x_j, \cdots)_1$. Now, $(\cdots, b_i x_j, \cdots)_2 \supseteq (\cdots, b_i x_j, \cdots)_1 = (a_1, \cdots, a_t) \supseteq a_i$, whence $\mathfrak{b}_2\mathfrak{x}_2 \supseteq (a_1, \cdots, a_t)_2 = \mathfrak{a}_2$. On the other hand, $(a_1, \cdots, a_t)_2 \supseteq (a_1, \cdots, a_t)_1 \supseteq b_i x_j$, so that $\mathfrak{a}_2 \supseteq (\cdots, b_i x_j, \cdots)_2 = \mathfrak{b}_2\mathfrak{x}_2$. Hence $\mathfrak{a}_2 = \mathfrak{b}_2\mathfrak{x}_2$.

THEOREM 8. *Let the hypotheses of Theorem 7 hold. Then if $\mathfrak{M}_1$ has property A so also has $\mathfrak{M}_2$.*

If $\mathfrak{a}_1 = (d)_1 = \mathfrak{o}_1(d)_1$, it follows by Theorem 7 that $\mathfrak{a}_2 = \mathfrak{o}_2(d)_2 = (d)_2$. Since every $\mathfrak{a}_2$ is associated with an $\mathfrak{a}_1$, it follows that $\mathfrak{M}_2$ has property A.

THEOREM 9. *Let the hypotheses of Theorem 7 hold. Then if $\mathfrak{M}_1$ has property B, so also has $\mathfrak{M}_2$.*

Let $\mathfrak{a}_2 \subseteq \mathfrak{b}_{2_s}$. Now, it must be noted that although $\mathfrak{b}_1 \subseteq \mathfrak{b}_2$, it is true that $\mathfrak{b}_{1_s} = \mathfrak{b}_{2_s}$, since $\mathfrak{b}_1$ and $\mathfrak{b}_2$ are determined by the same set of basal elements $b_1, \cdots, b_n$. Hence $\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \mathfrak{b}_{2_s} = \mathfrak{b}_{1_s}$, whence $\mathfrak{a}_1 = \mathfrak{b}_1\mathfrak{x}_1$, where $\mathfrak{x}_1$ is an ideal in $\mathfrak{M}_1$. Hence, by Theorem 7, $\mathfrak{a}_2 = \mathfrak{b}_2\mathfrak{x}_2$ and $\mathfrak{M}_2 B$ holds.

It must be remarked that the property C is also monotonic, as $\mathfrak{a}_1\mathfrak{b}_1 = (\gamma)_1$ implies $\mathfrak{a}_2\mathfrak{b}_2 = (\gamma)_2$.

That property D is not monotonic will be seen later in this paper.

THEOREM 10. *Let the hypotheses of Theorem 7 hold. Then if $\mathfrak{M}_2$ has property E, so also has $\mathfrak{M}_1$.*

If $(a)_1\mathfrak{c}_1 \subseteq (b)_1\mathfrak{c}_1$, then if $\mathfrak{c}_1 = (c_1, \cdots, c_m)_1$, we have $ac_i \subseteq (bc_1, \cdots, bc_m)_1 \subseteq (bc_1, \cdots, bc_m)_2 = (b)_2\mathfrak{c}_2$, whence $(a)_2\mathfrak{c}_2 \subseteq (b)_2\mathfrak{c}_2$ and $a = bm$, where m is in $\mathfrak{o}$.

5. **The ideal system $\mathfrak{L}$.**[6] We define an ideal system $\mathfrak{L}$ by letting $(a_1, \cdots, a_n)$ be the totality of all $\sum_{i=1}^{n} \xi_i a_i$, where ξ_i is in $\mathfrak{o}$. That $\mathfrak{L}$ actually is an ideal system is easily seen, as the necessary properties (1)–(5) for ideals are satisfied.

We say that $\mathfrak{R}$ is *$\mathfrak{o}$-complete* (*relative to* $\mathfrak{g}$) if for every two elements a and b in $\mathfrak{g}$ there is a greatest common $\mathfrak{o}$-divisor d so that $a = d\xi$, $b = d\eta$, where ξ and η are in $\mathfrak{o}$. If, in addition, $d = \lambda a + \mu b$, where λ and μ are also in $\mathfrak{o}$, $\mathfrak{R}$ is *linearly $\mathfrak{o}$-complete* (*relative to* $\mathfrak{g}$).

By $\mathfrak{L}A$ we indicate that ideal system $\mathfrak{L}$ has property A—similarly for $\mathfrak{L}B$, etc.

THEOREM 11. *$\mathfrak{R}$ is linearly $\mathfrak{o}$-complete relative to $\mathfrak{g}$ if and only if it has property $\mathfrak{L}A$.*

For, if $(a_1, \cdots, a_n) = (d)$, then $a_1, \cdots, a_n$ have d as a greatest common $\mathfrak{o}$-divisor and also $d = \sum_{i=1}^{n} \xi_i a_i$, ξ_i in $\mathfrak{o}$. Conversely, if $a_1, \cdots, a_n$ have d as a greatest common $\mathfrak{o}$-divisor, then $(a_1, \cdots, a_n) \subseteq (d)$. But since $d = \sum \xi_i a_i$, it follows that $(d) \subseteq (a_1, \cdots, a_n)$, whence $(a_1, \cdots, a_n) = (d)$.

THEOREM 12. *Properties $\mathfrak{L}B$ and $\mathfrak{L}C$ are equivalent. In fact, property $\mathfrak{L}C$ and $\mathfrak{a} \subseteq \mathfrak{b}$ for ideals $\mathfrak{a}$ and $\mathfrak{b}$ in $\mathfrak{L}$ are sufficient to prove that $\mathfrak{a}$ is divisible by $\mathfrak{b}$.*

The proof of this theorem can be made as in the theory of algebraic numbers. See, for example, Hecke, *Theorie der algebraischen Zahlen*, p. 93.

THEOREM 13. *$\mathfrak{L}C$ obtains if and only if for every set of integral elements $a_1, \cdots, a_n$ there exist elements $\xi_1, \cdots, \xi_n$ in $\mathfrak{o}$ such that*

$$a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n = a_1 \tag{1}$$

and such that $a_i\xi_j$ $(i, j = 1, \cdots, n)$ is in (a_1).

The conditions are sufficient. For,

$$\begin{aligned}(a_1, \cdots, a_n)(\xi_1, \cdots, \xi_n) &= (a_1\xi_1, \cdots, a_1\xi_n, a_2\xi_1, \cdots, a_n\xi_n) \\ &= (a_1, a_2\xi_1, a_2\xi_2, \cdots, a_n\xi_n) = (a_1),\end{aligned}$$

whence $\mathfrak{L}C$ holds.

Conversely, let $\mathfrak{L}C$ hold. Then, by Theorem 12, $\mathfrak{L}B$ holds, also. We have

$$(a_1) \subseteq (a_1, \cdots, a_n),$$

whence, by Theorem 12, $(a_1) = (a_1, \cdots, a_n)(b_1, \cdots, b_m)$. Therefore, $a_i b_j = a_1 y_{ij}$ with y_{ij} in $\mathfrak{o}$; in particular, for $i = 1$, $a_1 b_j = a_1 y_{1j}$, whence $b_j = y_{1j}$ is in $\mathfrak{o}$, for $j = 1, \cdots, m$. Hence $a_i b_j$ is in (a_1) for $i = 1, \cdots, n$; $j = 1, \cdots, m$.

[6] Prüfer, op. cit., p. 10.

We have, then, $a_1 = \sum a_i b_j x_{ij}$ (x_{ij} in $\mathfrak{o}$) so that if we write $\xi_i = \sum_{j=1}^{m} b_j x_{ij}$ (whence ξ_i is in $\mathfrak{o}$ and $a_i\xi_j$ is in (a_1)), then

$$a_1 = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n .$$

THEOREM 14. *Properties $\mathfrak{L}C$ and $\mathfrak{L}D$ are equivalent.*

We first prove the

LEMMA. *In ideal system $\mathfrak{L}$, property D implies $(ab) \subseteq (a^2, b^2)$, whence $(a^2, b^2) = (a, b)^2$.*

We have

$$(ab)(a, b) = (a^2b, ab^2) \subseteq (a^3, ab^2, a^2b, b^3) = (a^2, b^2)(a, b)$$

whence, by property D,

$$(ab) \subseteq (a^2, b^2).$$

Hence $(a, b)^2 = (a^2, ab, b^2) = (a^2, b^2)$.

We now proceed to the proof of the theorem and show first that it holds for all ideals in $\mathfrak{L}$ that are generated by two elements; we then complete the proof by induction on the number of generating elements.

Consider the ideal (a, b). By the lemma, there exist elements X and Y in $\mathfrak{o}$ such that

$$ab = Xa^2 + Yb^2.$$

Hence

$$\begin{aligned}(X)(a^2, b^2) &= (Xa^2, Xb^2)\\ &= (ab - Yb^2, Xb^2)\\ &\subseteq (ab, b^2) = (b)(a, b),\end{aligned}$$

whence by the lemma and property D,

$$(X)(a, b) \subseteq (b).$$

Hence $Xa = bX'$ (X' in $\mathfrak{o}$) whence $a = X'a + Yb$, with $X'b$ in (a). Hence (as in the proof of Theorem 13),

$$\begin{aligned}(a, b)(X', Y) &= (aX', bX', aY, bY)\\ &= (a, bX', bY) = (a)\end{aligned}$$

and the theorem is proved for all ideals in $\mathfrak{L}$ that are generated by two elements.

We now show that every ideal that is generated by $n + 1$ elements can be multiplied by a suitable ideal to yield an ideal generated by n elements, so that the theorem will be proved. Hence we now assume, for purposes of the induction, that the theorem is true for every ideal in $\mathfrak{L}$ that is generated by n elements.

Consider the ideal $(a, a_1, \cdots, a_{n-1})$; then, by Theorem 13, there exist elements X, Y, ξ_i $(i = 0, 1, \cdots, n-1)$ such that

$$a = \xi_0 a + \sum_{i=1}^{n-1} \xi_i a_i, \qquad a = Xa + Ya_n,$$

where $\xi_i a_j = m_{ij} a$, $Xa_n = ma$ $(i, j = 1, \cdots, n-1)$ and m_{ij}, m are in $\mathfrak{o}$. Hence

$$\begin{aligned} a &= \xi_0(Xa + Ya_n) + \sum_{i=1}^{n-1} \xi_i a_i \\ &= \xi_0 Xa + \sum_{i=1}^{n-1} \xi_i a_i + \xi_0 Y a_n . \end{aligned}$$

Now, we have

$$\begin{aligned} &(a, a_1, \cdots, a_{n-1}, a_n)(\xi_0 X, \xi_1, \xi_2, \cdots, \xi_{n-1}, \xi_0 Y) \\ &\qquad = (a\xi_0 X, \cdots, a\xi_{n-1}, a\xi_0 Y, \cdots, a_i \xi_0 X, \cdots, a_i \xi_j, \cdots, a_i \xi_0 Y, \cdots) \\ &\qquad = (a, a_n \xi_1, a_n \xi_2, \cdots, a_n \xi_{n-1}), \end{aligned}$$

using the conditions above. But the product ideal is generated by n elements and the theorem is proved.

Theorem 15. *Property $\mathfrak{P}E$ holds if and only if every equation of the form*

$$a^n + \xi_1 a^{n-1} b + \xi_2 a^{n-2} b^2 + \cdots + \xi_n b^n = 0, \tag{2}$$

where ξ_i is in $\mathfrak{o}$, implies that $a = b\xi$, ξ in $\mathfrak{o}$.

Let $\mathfrak{P}E$ hold. Then if an equation (2) holds, we have

$$\begin{aligned} &(a)(b^{n-1}, b^{n-2}a, \cdots, ba^{n-2}, a^{n-1}) \\ &\qquad = (ab^{n-1}, a^2 b^{n-2}, \cdots, a^{n-1} b, -\xi_1 a^{n-1} b - \xi_2 a^{n-2} b^2 - \cdots - \xi_n b^n) \\ &\qquad = (ab^{n-1}, a^2 b^{n-2}, \cdots, -\xi_n b^n) \subseteq (b)(ab^{n-2}, a^2 b^{n-3}, \cdots, a^{n-1}, b^{n-1}), \end{aligned}$$

whence $a = b\xi$, ξ in $\mathfrak{o}$.

Conversely, let every equation of the form (2) hold only if $a = b\xi$, with ξ in $\mathfrak{o}$. If

$$(a)(c_1, \cdots, c_m) \subseteq (b)(c_1, \cdots, c_m),$$

then

$$ac_i = \sum_{j=1}^{m} \lambda_{ij} b c_j \qquad (i = 1, \cdots, m),$$

where λ_{ij} is in $\mathfrak{o}$. Hence if $\delta_{ii} = 1$, $\delta_{ij} = 0$ for $i \neq j$,

$$|\lambda_{ij} b - \delta_{ij} a| = 0,$$

an equation of the form (2), whence $a = b\xi$, with ξ in $\mathfrak{o}$, and $\mathfrak{P}E$ obtains.

It will be seen later that $\mathfrak{P}D$ and $\mathfrak{P}E$ are inequivalent properties.

THEOREM 16. *Let $\mathfrak{R}$ have property $\mathfrak{L}C$ relative to $\mathfrak{g}$ and $\mathfrak{o}$. If every pair of the finite set of congruences*

$$(3) \qquad x \equiv r \pmod{a}, \qquad x \equiv s \pmod{b}, \qquad x \equiv t \pmod{c}, \qquad \cdots$$

($a, b, c, \cdots, r, s, t, \cdots$ in $\mathfrak{g}$) is solvable, then there exists an element ξ in $\mathfrak{o}$ such that the congruences

$$(4) \qquad x \equiv \xi r \pmod{a}, \qquad x \equiv \xi s \pmod{b}, \qquad x \equiv \xi t \pmod{c}, \qquad \cdots$$

are simultaneously solvable.

We prove the theorem by induction. We consider the congruences obtained from (3) by omitting first the first congruence and then the second. We obtain, then, by the hypothesis of the induction, two elements ξ_1 and ξ_2 in $\mathfrak{o}$ such that the congruences

$$x \equiv \xi_1 s \pmod{b}, \qquad x \equiv \xi_1 t \pmod{c}, \qquad \cdots$$

and

$$y \equiv \xi_2 r \pmod{a}, \qquad y \equiv \xi_2 t \pmod{c}, \qquad \cdots$$

have simultaneous solutions x and y.

Since $\mathfrak{L}C$ obtains, there exist, by Theorem 13, elements ξ and η in $\mathfrak{o}$ such that

$$\xi a + \eta(\xi_2 x - \xi_1 y) = a,$$

where $\xi(\xi_2 x - \xi_1 y)$ is in (a). Then

$$\xi\xi_2 x \equiv \xi\xi_1\xi_2 s \pmod{b}, \qquad \xi\xi_2 x \equiv \xi\xi_1\xi_2 t \pmod{c}, \qquad \cdots,$$

$$\xi\xi_1 y \equiv \xi\xi_1\xi_2 r \pmod{a}, \qquad \xi\xi_1 y \equiv \xi\xi_1\xi_2 t \pmod{c}, \qquad \cdots,$$

and $\xi\xi_2 x \equiv \xi\xi_1 y \pmod{a}$. Hence we have a simultaneous solution of (4) with ξ replaced by $\xi\xi_1\xi_2$ and x by $\xi\xi_2 x$.

COROLLARY. *If $\mathfrak{L}C$ obtains, there exists an element ξ in $\mathfrak{o}$ such that every integer in (a, b) which is a multiple of ξ can be written as*

$$m(a + b) + \textit{common multiple of } a \textit{ and } b.$$

For, every pair of the congruences

$$x \equiv 0 \pmod{a}, \qquad x \equiv 0 \pmod{b}, \qquad x \equiv Xa + Yb \pmod{a + b}$$

is solvable.

6. **The ideal system $\mathfrak{A}$.** Following Prüfer, we state the following definition: *Let $(a_1, \cdots, a_n)_{\mathfrak{A}}$ be the set of elements a in $\mathfrak{g}$ for which there exists an ideal $\mathfrak{r}_{\mathfrak{L}} \neq (0)$ in $\mathfrak{L}$ such that*

$$(a)\mathfrak{r}_{\mathfrak{L}} \subseteq (a_1, \cdots, a_n)_{\mathfrak{L}}\mathfrak{r}_{\mathfrak{L}}.$$

If the ring $\mathfrak{R}$ is integrally closed relative to $\mathfrak{g}$ and $\mathfrak{o}$ (that is, $\mathfrak{L}E$ obtains), this set $(a_1, \cdots, a_n)_{\mathfrak{A}}$ is an ideal in an ideal system which will be denoted by $\mathfrak{A}$.

The proof of the ideal property can be found in Prüfer's paper.[7] We note that always $(a_1, \cdots, a_n)_\mathfrak{L} \subseteq (a_1, \cdots, a_n)_\mathfrak{A}$.

Prüfer also proves[7] the following two theorems and we merely state them.

THEOREM 17. *If* $a_1, \cdots, a_m$ *are contained in* $(b_1, \cdots, b_n)_\mathfrak{A}$, *then there exists an ideal* $\mathfrak{x}_\mathfrak{L} \neq (0)$ *in* $\mathfrak{L}$ *such that*

$$(a_1, \cdots, a_m)_\mathfrak{L}\, \mathfrak{x}_\mathfrak{L} \subseteq (b_1, \cdots, b_n)_\mathfrak{L}\, \mathfrak{x}_\mathfrak{L}.$$

THEOREM 18. *An element* a *in* $\mathfrak{g}$ *belongs to* $(a_1, \cdots, a_n)_\mathfrak{A}$ *if and only if* a *satisfies an equation*

$$a^k + f_1(a_1, \cdots, a_n)a^{k-1} + \cdots + f_k(a_1, \cdots, a_n) = 0,$$

where $f_i(a_1, \cdots, a_n)$ *is a homogeneous polynomial of degree* i *in* $a_1, \cdots, a_n$, *with coefficients in* $\mathfrak{o}$.

THEOREM 19. $\mathfrak{A}A$ *and* $\mathfrak{L}A$ *are equivalent properties.*

$\mathfrak{L}A$ implies $\mathfrak{A}A$ by Theorem 8. Conversely, let $\mathfrak{A}A$ hold so that $(a_1, \cdots, a_n) = (d)$. Then $a_i = d\mu_i$, where μ_i is in $\mathfrak{o}$. Also, by Theorem 18, d satisfies an equation

$$d^k + f_1 d^{k-1} + f_2 d^{k-2} + \cdots + f_k = 0,$$

where f_i is a homogeneous polynomial of degree i in $a_1, \cdots, a_n$, with coefficients in $\mathfrak{o}$. Hence $f_i = d^{i-1} f_i'$, where f_i' is a *linear* form in $a_1, \cdots, a_n$, with coefficients in $\mathfrak{o}$, so that

$$d^k + d^{k-1} \sum f_i' = 0$$

and

$$d = -\sum f_i',$$

a linear form in $a_1, \cdots, a_n$, with coefficients in $\mathfrak{o}$. By Theorem 11, $\mathfrak{L}A$ holds.

THEOREM 20. $\mathfrak{A}C$ *and* $\mathfrak{L}C$ *are equivalent properties.*

$\mathfrak{L}C$ implies $\mathfrak{A}C$ by the remark following the proof of Theorem 9. Conversely, let $\mathfrak{A}C$ hold. If $\mathfrak{b}$ is an ideal in $\mathfrak{A}$, then there exists $\mathfrak{c}$ such that $\mathfrak{bc} = (\gamma)$, or if $\mathfrak{b} = (b_1, \cdots, b_s)$, $\mathfrak{c} = (c_1, \cdots, c_t)$, $(\gamma) = (\cdots, b_i c_j, \cdots)$. Hence

$$\gamma^k + f_1 \gamma^{k-1} + f_2 \gamma^{k-2} + \cdots + f_k = 0,$$

where f_i is a homogeneous polynomial of degree i in $b_1 c_1, \cdots, b_s c_t$, with coefficients in $\mathfrak{o}$. Hence $f_i = \gamma^{i-1} f_i'$, where f_i' is *linear* in $b_1 c_1, \cdots, b_s c_t$, with coefficients in $\mathfrak{o}$, so that we have

$$\gamma^k + (f_1' + f_2' + \cdots + f_k')\gamma^{k-1} = 0,$$

or $\gamma = \sum \lambda_{ij} b_i c_j$, where λ_{ij} is in $\mathfrak{o}$. But since $(b_i c_j)_\mathfrak{L} \subseteq (\gamma)_\mathfrak{L}$ it follows that $(\cdots, b_i c_j, \cdots)_\mathfrak{L} = \mathfrak{b}_\mathfrak{L} \mathfrak{c}_\mathfrak{L} = (\gamma)_\mathfrak{L}$. Hence $\mathfrak{L}C$ obtains.

COROLLARY. $\mathfrak{A}B$, $\mathfrak{A}C$, $\mathfrak{L}B$, $\mathfrak{L}C$ *and* $\mathfrak{L}D$ *are equivalent properties.*

For, by Theorems 12 and 14, $\mathfrak{L}B$, $\mathfrak{L}C$ and $\mathfrak{L}D$ are equivalent properties.

[7] Pp. 14–15.

THEOREM 21. *$\mathfrak{L}E$ implies $\mathfrak{A}D$, that is to say, property D always holds in ideal system $\mathfrak{A}$.*[8]

If

$$(a_1, \cdots, a_m)(c_1, \cdots, c_s) \subseteq (b_1, \cdots, b_n)(c_1, \cdots, c_s),$$

then Theorem 17 yields the existence of an ideal $\mathfrak{x}_{\mathfrak{L}} \neq (0)$ such that

$$(a_1, \cdots, a_m)_{\mathfrak{L}}(c_1, \cdots, c_s)_{\mathfrak{L}}\mathfrak{x}_{\mathfrak{L}} \subseteq (b_1, \cdots, b_n)_{\mathfrak{L}}(c_1, \cdots, c_s)_{\mathfrak{L}}\mathfrak{x}_{\mathfrak{L}}.$$

Since $(c_1, \cdots, c_s) \neq (0) \neq \mathfrak{x}_{\mathfrak{L}}$, $(c_1, \cdots, c_s)_{\mathfrak{L}}\mathfrak{x}_{\mathfrak{L}} = \mathfrak{U}_{\mathfrak{L}} \neq (0)$ and

$$(a_1, \cdots, a_m)_{\mathfrak{L}}\mathfrak{U}_{\mathfrak{L}} \subseteq (b_1, \cdots, b_n)_{\mathfrak{L}}\mathfrak{U}_{\mathfrak{L}}$$

and by definition of ideal system $\mathfrak{A}$, $(a_1, \cdots, a_m) \subseteq (b_1, \cdots, b_n)$, whence $\mathfrak{A}D$ holds.

COROLLARY. *$\mathfrak{L}E$, $\mathfrak{A}D$, and $\mathfrak{A}E$ are equivalent properties.*

For, $\mathfrak{L}E$ implies $\mathfrak{A}D$ by the preceding theorem, $\mathfrak{A}D$ implies $\mathfrak{A}E$ always, and $\mathfrak{A}E$ implies $\mathfrak{L}E$ by Theorem 10.

7. **The ideal system $\mathfrak{U}$.** If a and b are elements in $\mathfrak{g}$, then b is said to be an $\mathfrak{o}$-*divisor of* a (a is an $\mathfrak{o}$-*multiple of* b) if $a = bc$, where c is in $\mathfrak{o}$.

Let $(a_1, \cdots, a_n)_{\mathfrak{U}}$ represent the totality of elements of $\mathfrak{g}$ which are $\mathfrak{o}$-multiples of all common $\mathfrak{o}$-divisors of $a_1, \cdots, a_n$. $(a_1, \cdots, a_n)_{\mathfrak{U}}$ is an ideal in $\mathfrak{R}$ relative to $\mathfrak{g}$ and $\mathfrak{o}$ and the set of all such is an ideal system $\mathfrak{U}$.

We note that $\mathfrak{a} = \mathfrak{a}_{\mathfrak{U}}$ may be (0) even if the generating elements are not all zero. For, if $\mathfrak{o}$ is a proper subring of $\mathfrak{g}$, it can very well happen that two elements in $\mathfrak{g}$ may have no common $\mathfrak{o}$-divisors so that the ideal in $\mathfrak{U}$ generated by them is the zero ideal.

We have to verify properties (1)–(5) for an ideal.

1. Every a_i is an $\mathfrak{o}$-multiple of every common $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$.

2. Let $a_1, \cdots, a_m$ be in $(b_1, \cdots, b_n)$. Then if d is a common $\mathfrak{o}$-divisor of $b_1, \cdots, b_n$, it follows that $a_1, \cdots, a_m$ are $\mathfrak{o}$-multiples of d, whence, since d is *any* common $\mathfrak{o}$-divisor of $b_1, \cdots, b_n$,

$$(a_1, \cdots, a_m)_{\mathfrak{U}} \subseteq (b_1, \cdots, b_n)_{\mathfrak{U}}.$$

3. If a_1 is in (a), then a_1 is an $\mathfrak{o}$-multiple of a, since the $\mathfrak{o}$-multiples of all $\mathfrak{o}$-divisors of a are exactly the $\mathfrak{o}$-multiples of a.

4. If a is in $(a_1, \cdots, a_n)$, then ab is a product of b and an $\mathfrak{o}$-multiple of all common $\mathfrak{o}$-divisors of $a_1, \cdots, a_n$, so that ab is in $(a_1b, \cdots, a_nb)$.

5. The element $a + b$ is an $\mathfrak{o}$-multiple of every common $\mathfrak{o}$-divisor of a and b.

THEOREM 22. *The ideal $(a_1, \cdots, a_n)$ in any ideal system is a subset of $(a_1, \cdots, a_n)_{\mathfrak{U}}$, provided that $(a_1, \cdots, a_n)_{\mathfrak{U}} \neq (0)$.*

For, if d is any common $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$, then $(a_1, \cdots, a_n) \subseteq (d)$ whence $(a_1, \cdots, a_n)$ can contain only $\mathfrak{o}$-multiples of d. Since d is *any* common

[8] For ideal system $\mathfrak{A}$, property $\mathfrak{L}E$ is assumed always to hold.

$\mathfrak{o}$-divisor of $a_1, \cdots, a_n$, it follows that $(a_1, \cdots, a_n)$ contains only $\mathfrak{o}$-multiples of all common $\mathfrak{o}$-divisors of $a_1, \cdots, a_n$. But then, by definition, $(a_1, \cdots, a_n) \subseteq (a_1, \cdots, a_n)_{\mathfrak{U}}$.

THEOREM 23. *$\mathfrak{R}$ is $\mathfrak{o}$-complete if and only if $\mathfrak{U}A$ holds.*

If $\mathfrak{R}$ is $\mathfrak{o}$-complete and if $(a_1, \cdots, a_n)_{\mathfrak{U}} \neq (0)$, we let d be the greatest common $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$. Then $(d) \supseteqq (a_1, \cdots, a_n)_{\mathfrak{U}}$ since every $\mathfrak{o}$-multiple of all common $\mathfrak{o}$-divisors of $a_1, \cdots, a_n$ is an $\mathfrak{o}$-multiple of d. But $(d) \subseteq (a_1, \cdots, a_n)_{\mathfrak{U}}$ since d must be $\mathfrak{o}$-divisible by every $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$. Hence $(a_1, \cdots, a_n)_{\mathfrak{U}} = (d)$.

Conversely, let $\mathfrak{U}A$ hold, $(a_1, \cdots, a_n)_{\mathfrak{U}} = (d) \neq (0)$. Then d must be $\mathfrak{o}$-divisible by every $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$; but d itself is an $\mathfrak{o}$-divisor of the a_i. Hence d is the greatest common $\mathfrak{o}$-divisor of $a_1, \cdots, a_n$ and $\mathfrak{R}$ is $\mathfrak{o}$-complete relative to $\mathfrak{g}$ and $\mathfrak{o}$.

$\mathfrak{U}B$ and $\mathfrak{U}A$ are inequivalent properties for we need consider only the case when $\mathfrak{R}$ is an (absolute) algebraic number field and $\mathfrak{g} = \mathfrak{o}$ is the set of all integers in $\mathfrak{R}$. Then $\mathfrak{U}B$ holds, but if $\mathfrak{R}$ has class number >1, $\mathfrak{U}A$ does not hold.

On the other hand, $\mathfrak{U}B$ and $\mathfrak{U}C$ are equivalent properties and the proof can be made in the same way as that of $\mathfrak{L}B$ and $\mathfrak{L}C$.

THEOREM 24. *$\mathfrak{U}C$ and $\mathfrak{U}D$ are inequivalent properties.*

We first find a necessary condition that property $\mathfrak{U}C$ obtain. Let $\mathfrak{a} = (a_1, \cdots, a_p)$, $\mathfrak{b} = (b_1, \cdots, b_r)$ and $\mathfrak{a}\mathfrak{b} = (\gamma)$. Then there exist pr quantities m_{ij} in $\mathfrak{o}$ such that

$$a_i b_j = \gamma m_{ij}.$$

At once we find that $a_i m_{jk} = a_j m_{ik}$. We write $m_{i1} = m_i$. Then

$$\begin{aligned}(a_1, \cdots, a_p)&(m_1, \cdots, m_{p-1})\\ &= (a_1m_1, \cdots, a_1m_{p-1}, a_2m_1, \cdots, a_2m_{p-1}, \cdots, a_pm_1, \cdots, a_pm_{p-1})\\ &= (a_1m_1, \cdots, a_1m_{p-1}, a_1m_2, \cdots, a_{p-1}m_2, \cdots, a_1m_p, \cdots, a_{p-1}m_p)\\ &= (a_1, \cdots, a_{p-1})(m_1, \cdots, m_{p-1}, m_p).\end{aligned}$$

Hence, if $\mathfrak{a}$ is any ideal and a any element in $\mathfrak{g}$, there exist an ideal $\mathfrak{m}$ in $\mathfrak{o}$ and an element m in $\mathfrak{o}$ such that

$$(\mathfrak{a}, a)\mathfrak{m} = (\mathfrak{m}, m)\mathfrak{a}. \tag{5}$$

Hence, if $\mathfrak{U}C$ is to hold, an equation (5) must hold for all choices of $\mathfrak{a}$ and a.

We now adduce an example in which (5) cannot hold for all $\mathfrak{a}$ and a so that $\mathfrak{U}C$ cannot obtain; property $\mathfrak{U}D$, however, can be shown to hold.

We take $\mathfrak{R} = \mathfrak{g}$ to be a transcendental ring extension of a finite field $\mathfrak{o}$ of q elements, where q is a rational prime. Since every ideal in $\mathfrak{U}$ which is generated by elements of $\mathfrak{o}$ must be $\mathfrak{o}$ itself, the condition (5) becomes

$$(\mathfrak{a}, a) = \mathfrak{a},$$

a condition which is certainly not satisfied if a is not in $\mathfrak{a}$ and $\mathfrak{a} \neq \mathfrak{g}$ (all of this is possible since $\mathfrak{g}$ is not a field).

It remains to show that $\mathfrak{U}D$ holds.

If $\mathfrak{a} = (a_1, \cdots, a_p)$, $\mathfrak{b} = (b_1, \cdots, b_r)$, $\mathfrak{c} = (c_1, \cdots, c_s)$, and if $\mathfrak{a}\mathfrak{b} \subseteq \mathfrak{a}\mathfrak{c}$, then

$$a_i b_j \subseteq (\cdots, a_k c_l, \cdots).$$

Hence, by the definition of the ideal system $\mathfrak{U}$ there exist, for every $\mathfrak{o}$-divisor δ of $a_i c_k$ ($i = 1, \cdots, p$; $k = 1, \cdots, s$), elements m_{ij} and n_{ij} in $\mathfrak{o}$ such that

$$a_i b_j = \delta m_{ij}, \qquad a_i c_k = \delta n_{ik}.$$

Since $\alpha^{q-1} = 1$ for all elements $\alpha \neq 0$ in $\mathfrak{o}$, we have

$$a_i^{q-1} b_j^{q-1} = \delta^{q-1} = a_i^{q-1} c_k^{q-1},$$

whence $b_j^{q-1} = c_k^{q-1}$. Hence $b_j = c_k \xi_{jk}$, where $\xi_{jk}^{q-1} = 1$, so that ξ_{jk} must be one of $1, 2, \cdots, q - 1$ and is, therefore, in $\mathfrak{o}$. Hence $(b_j) \subseteq (c_k)$ for all j and k so that $\mathfrak{b} \subseteq \mathfrak{c}$.

THEOREM 25. *$\mathfrak{U}D$ and $\mathfrak{U}E$ are inequivalent properties.*

We adduce the following example. Let $\mathfrak{R}$ be the field of rational functions of two indeterminates ξ and η over the field of ordinary rational numbers. We take $\mathfrak{g} = \mathfrak{o}$ to be the set of functions of $\mathfrak{R}$ in which (1) the total degree of ξ and η in the numerator is less than or equal to the total degree of ξ and η in the denominator, and (2) the degree of ξ in the numerator is less than or equal to the degree of ξ in the denominator. It is easily seen that $\mathfrak{g}$ is a subring of $\mathfrak{R}$.

$\mathfrak{U}D$ cannot hold. For,

$$\begin{aligned}\left(\frac{1}{\xi}, \frac{1}{\eta}\right)\left(\frac{1}{\xi}, \frac{1}{\eta^2}\right) &= \left(\frac{1}{\xi^2}, \frac{1}{\xi\eta}, \frac{1}{\xi\eta^2}, \frac{1}{\eta^3}\right)\\ &= \left(\frac{1}{\xi\eta}\cdot\frac{\eta}{\xi}, \frac{1}{\xi\eta}, \frac{1}{\xi\eta^2}, \frac{1}{\eta^3}\right)\\ &= \left(\frac{1}{\xi\eta}, \frac{1}{\xi\eta^2}, \frac{1}{\eta^3}\right)\\ &= \left(\frac{1}{\eta}\right)\left(\frac{\eta}{\xi\eta}, \frac{1}{\xi\eta}, \frac{1}{\eta^2}\right)\\ &= \left(\frac{1}{\eta}\right)\left(\frac{1}{\xi\eta}, \frac{1}{\eta^2}\right)\\ &= \left(\frac{1}{\eta^2}\right)\left(\frac{1}{\xi}, \frac{1}{\eta}\right),\end{aligned}$$

whence, if $\mathfrak{U}D$ is to hold,

$$\left(\frac{1}{\xi}, \frac{1}{\eta^2}\right) = \left(\frac{1}{\eta^2}\right). \tag{6}$$

But ξ^{-1} is not divisible by η^{-2}, for $\eta^2\xi^{-1}$ is not in $\mathfrak{g}$. Hence (6) cannot hold and $\mathfrak{U}D$ does not obtain.

We now show that $\mathfrak{U}E$ holds. Let $(a)(a_1, \cdots, a_m) \subseteq (b)(a_1, \cdots, a_m)$. We may write $a_1 = \eta^{-\lambda}(\eta^{\lambda}a_1) = \eta^{-\lambda}a_1'$, where the total degree of the numerator of a_1' is the same as that of its denominator. Again, $a_1' = \xi^{-\mu}(\xi^{\mu}a_1') = \xi^{-\mu}a_1''$, where a_1'' has, in addition to the property possessed by a_1', the additional one of having the ξ-degree of its numerator equal to that of its denominator. Hence $a_1 = \xi^{-\mu}\eta^{-\lambda}a_1''$, where a_1'' and $1/a_1''$ are in $\mathfrak{g}$. It may not be so that for every i, $1 \leqq i \leqq m$, $a_i = \xi^{-\mu}\eta^{-\lambda}a_i''$ with a_i'' in $\mathfrak{g}$, but corresponding to some a_k there is a pair of integers λ' and μ' such that (1) a_i is divisible by $\delta' = \xi^{-\mu'}\eta^{-\lambda'}$ for $1 \leqq i \leqq m$, and (2) $a_k = \delta' a_k''$, where a_k'' and $1/a_k''$ are in $\mathfrak{g}$.

Now, since $aa_k = b\delta' m$, where m is in $\mathfrak{g}$, we have $aa_k'' = bm$, whence $a = b(m/a_k'')$ so that, since m/a_k'' is in $\mathfrak{g}$, $(a) \subseteq (b)$ and $\mathfrak{U}E$ holds.

8. **The ideal systems $\mathfrak{L}$ and $\mathfrak{U}$. Extension**. Let $\mathfrak{R}$ be an arbitrary commutative ring with a unit element and possessing no divisors of zero. Let $\mathfrak{R}(z)$ be a transcendental extension of the quotient field of $\mathfrak{R}$ so that every element of $\mathfrak{R}(z)$ can be written as a quotient of polynomials in z with coefficients in $\mathfrak{R}$. Then a relation $(a_1, \cdots, a_m)_{\mathfrak{u}} \supseteqq (b_1, \cdots, b_n)_{\mathfrak{u}}$ in $\mathfrak{R}$ does not necessarily imply the same relation in $\mathfrak{R}(z)$.

We shall prove this statement by the exhibition of a suitable example. Accordingly, let $\mathfrak{R}$ be the field of rational functions with ordinary rational coefficients of two indeterminates x and y; we define the set of integers $\mathfrak{g}$ to be those elements in which the total degree of x and y in the numerator does not exceed the corresponding degree in the denominator. $\mathfrak{g}$ is obviously a ring with unit and we take $\mathfrak{g} = \mathfrak{o}$.

We shall introduce a simple but convenient notation. If $\alpha = n/d$, where n and d are polynomials in x and y, then $D_{n(\alpha)}(x)$ represents the degree of x in n and $D_{n(\alpha)}(x, y)$ the degree of x and y in n—similarly for $D_{d(\alpha)}(x)$, $D_{d(\alpha)}(x, y)$.

We consider the ideal $(a, a')_{\mathfrak{u}}$ in $\mathfrak{R}$, where $a = x^{-2}$, $a' = y^{-2}$ are in $\mathfrak{g}$. It is easily seen that an integer φ divides a and a' if and only if $0 \leqq D_{d(\varphi)}(x, y) - D_{n(\varphi)}(x, y) \leqq 2$. Hence, if $b = xy^{-3}$, we have

$$D_{d(b\varphi^{-1})}(x, y) = 3 + D_{n(\varphi)}(x, y) \geqq D_{d(\varphi)}(x, y) + 1,$$

$$D_{n(b\varphi^{-1})}(x, y) = 1 + D_{d(\varphi)}(x, y),$$

so that $D_{d(b\varphi^{-1})}(x, y) \geqq D_{n(b\varphi^{-1})}(x, y)$ and b is divisible by φ. Since φ is an arbitrary common divisor of a and a', it follows that b is in the ideal $(a, a')_{\mathfrak{u}}$. We show now that with a properly chosen definition of integer in $\mathfrak{R}(z)$, b is *not* in $(a, a')_{\mathfrak{u}}$ for this larger ring.

In $\mathfrak{R}(z)$ we define the set of integers $\mathfrak{g}'$ to be those rational functions α of x, y, and z for which (1) $D_{d(\alpha)}(x, y) \geqq 1 + D_{n(\alpha)}(x, y)$, and (2) $D_{n(\alpha)}(x, z) \leqq D_{d(\alpha)}(x, z)$. That $\mathfrak{g}'$ is a subring with unit of $\mathfrak{R}(z)$ is clear and it is also clear that $\mathfrak{g} \subset \mathfrak{g}'$. We take $\mathfrak{o}' = \mathfrak{g}'$.

We can now show that in $\mathfrak{R}(z)$, with the notion of integral element just defined, $b = x/y^3$ is *not* in $(a, a')_{\mathfrak{u}}$. For, let $f = x/y^2z$; then f is in $\mathfrak{g}'$. Also $a/f = y^2z/x^3$, $a'/f = z/x$ so that f divides a and a'. But $b/f = z/y$ and is not in $\mathfrak{g}'$.

THEOREM 26. *Let $\mathfrak{R}$ be a commutative ring with unit element and possessing no divisors of zero and let $\mathfrak{g}$ be a subring of $\mathfrak{R}$ which has a unit but which is not a field. Then there exists in the transcendental extension $\mathfrak{R}(z)$ a subring $\mathfrak{g}' \supset \mathfrak{g}$ such that relative to $\mathfrak{g}'$ $\mathfrak{R}(z)$ does not have property $\mathfrak{L}C$.*

We define $\mathfrak{g}'$ to be the ring of all polynomials

$$\alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots ,$$

where α_0, α_1 are in $\mathfrak{g}$ and α_i $(i \geqq 2)$ are in $\mathfrak{R}$.

Let λ_0 in $\mathfrak{g}$ have the property that $\lambda_0 \neq 0$ and $\lambda_0\nu \neq 1$ for every ν in $\mathfrak{g}$—this can be realized since $\mathfrak{g}$ is not a field. Then, by Theorem 13, $\mathfrak{L}C$ can obtain in $\mathfrak{R}(z)$ only if there exist elements $X_0 + X_1 z + \cdots$, $Y_0 + Y_1 z + \cdots$ in $\mathfrak{g}'$ such that

$$\lambda_0(X_0 + X_1 z + \cdots) + \xi(Y_0 + Y_1 z + \cdots) = \lambda_0 ,$$

where $z(X_0 + X_1 z + \cdots) = \lambda_0(\nu_0 + \nu_1 z + \cdots)$, where $\nu_0 + \nu_1 z + \cdots$ is in $\mathfrak{g}'$. We have, immediately, $\lambda_0 X_0 = \lambda_0$, whence $X_0 = 1$; also $\lambda_0 \nu_1 = X_0 = 1$, contrary to the choice of λ_0. Hence $\mathfrak{L}C$ does not obtain.

COROLLARY 1. *There exist complete rings that are not linearly complete. That is to say, properties $\mathfrak{L}A$ and $\mathfrak{U}A$ are inequivalent.*

For, let $\mathfrak{R}$ be the ring of all rational integers and let $\mathfrak{R} = \mathfrak{g} = \mathfrak{o}$. Then, since $\mathfrak{R}$ is not a field, it follows from Theorem 26 that for a proper choice of the integral elements of $\mathfrak{R}(z)$, the property $\mathfrak{L}C$ does not hold in $\mathfrak{R}(z)$. Hence $\mathfrak{R}(z)$ cannot have property $\mathfrak{L}A$ although it does have property $\mathfrak{U}A$.

The example in the preceding paragraph yields the

COROLLARY 2. *Properties $\mathfrak{U}A$ and $\mathfrak{L}C$ are inequivalent.*

We mention, finally, that Prüfer has shown[9] the inequivalence of $\mathfrak{L}E$ and $\mathfrak{U}E$.

We now append the following scheme which shows the relations between the various ideal systems that we have considered. Implication is denoted by an arrow.

$$\begin{array}{ccccc}
\mathfrak{U}A & \leftarrow & \mathfrak{L}A & \leftrightarrow & \mathfrak{A}A \\
\downarrow & & \downarrow & & \downarrow \\
\mathfrak{U}B & \leftarrow & \mathfrak{L}B & \leftrightarrow & \mathfrak{A}B \\
\updownarrow & & \updownarrow & & \updownarrow \\
\mathfrak{U}C & \leftarrow & \mathfrak{L}C & \leftrightarrow & \mathfrak{A}C \\
\downarrow & & \updownarrow & & \downarrow \\
\mathfrak{U}D & \leftarrow & \mathfrak{L}D & \rightarrow & \mathfrak{A}D \\
\downarrow & & \downarrow & & \updownarrow \\
\mathfrak{U}E & \rightarrow & \mathfrak{L}E & \leftrightarrow & \mathfrak{A}E
\end{array}$$

A glance at this scheme reveals that $\mathfrak{A}D$ does not imply $\mathfrak{U}D$ so that D is not a monotonic property.

We prove, finally,

THEOREM 27. *If $\mathfrak{R}$ is a commutative ring with unit element having no divisors*

[9] Op. cit., p. 19.

of zero, then $\mathfrak{R}$ can be imbedded in a ring $\mathfrak{R}'$ which has no divisors of zero and such that

(1) *an element of $\mathfrak{R}$ is integral in $\mathfrak{R}'$ if and only if it is integral in $\mathfrak{R}$,*

(2) *$\mathfrak{R}'$ is complete.*

Define $\mathfrak{R}'$ to be the field of all rational functions with coefficients in the quotient field of $\mathfrak{R}$ of an indeterminate z. We define the set $\mathfrak{g}' = \mathfrak{o}'$ of integral elements of $\mathfrak{R}'$ to be the functions

$$\frac{a_0 + a_1 z + \cdots}{a_0' + a_1' z + \cdots},$$

where $a_0, a_1, \cdots, a_0', a_1', \cdots$ are in $\mathfrak{g} = \mathfrak{o}$, where a_0 is divisible by a_0', and where the degree of the numerator does not exceed that of the denominator. $\mathfrak{g}'$ is easily seen to be a subring of $\mathfrak{R}'$ with a unit element. Also, an element of $\mathfrak{R}$ can be integral in $\mathfrak{R}'$ if and only if it is integral in $\mathfrak{R}$.

Let α and β be in $\mathfrak{g}'$. We let d be a common denominator of α and β. Then if

$$D_{n(\alpha)}(z) = D_d(z) - m_1, \qquad D_{n(\beta)}(z) = D_d(z) - m_2,$$

where m_1 and m_2 are rational integers $\geqq 0$, we see that δ can divide both α and β only if

$$0 \leqq D_{d(\delta)}(z) - D_{n(\delta)}(z) \leqq \min(m_1, m_2).$$

But if we choose δ in particular to be

$$\delta_0 = \frac{1}{1 + z^{\min(m_1, m_2)}},$$

then δ_0 actually divides α and β and is in reality a greatest common divisor of α and β, since, as is easily verified, every common divisor of α and β must divide δ_0. Hence $\mathfrak{R}'$ is complete relative to $\mathfrak{g}'$.

YALE UNIVERSITY.

ONE-PARAMETER FAMILIES OF TRANSFORMATIONS

BY J. L. DOOB

One-parameter families of transformations taking measurable sets into measurable sets arise in many parts of mathematics. The purpose of this paper is to present a detailed study of the regularity properties of such families. Before we begin this study, some introductory remarks on measures in product spaces will be made. These remarks have some independent interest, so they will be phrased more generally than necessary for the actual applications to be made in the present paper.

Introduction

Let T (Ω) be any abstract space, whose points will be denoted by t (ω). Let $\mathcal{F}_t$ ($\mathcal{F}_\omega$) be any Borel field[1] of t-sets (ω-sets) including T (Ω) itself. The space of points (t, ω), the direct product of the two spaces, will be denoted by $T \times \Omega$. Let E be a set of $\mathcal{F}_t$ and let Λ be a set of $\mathcal{F}_\omega$. Then the condition $t \in E, \omega \in \Lambda$ determines a (t, ω)-set $E \times \Lambda$, and we shall denote by $\mathcal{F}_t \times \mathcal{F}_\omega$ the Borel field of (t, ω)-sets determined by all such sets $E \times \Lambda$.[2] If a Borel field of point sets is the Borel field determined by some denumerable subcollection, it will be called *strictly separable*. If $\mathcal{F}_t$ and $\mathcal{F}_\omega$ are strictly separable, $\mathcal{F}_t \times \mathcal{F}_\omega$ is also strictly separable. Moreover, if $\tilde{\Lambda}$ is any set of $\mathcal{F}_t \times \mathcal{F}_\omega$, there is a strictly separable subfield $\mathcal{F}'_t$ ($\mathcal{F}'_\omega$) of $\mathcal{F}_t$ ($\mathcal{F}_\omega$) such that $\tilde{\Lambda}$ is a set of the field $\mathcal{F}'_t \times \mathcal{F}'_\omega$.[3]

Let $\mu(\Lambda)$ be a non-negative completely additive set function defined on the field $\mathcal{F}_\omega$.[4] An ω-set Λ_1 will be called measurable if it differs from a set Λ_0 of $\mathcal{F}_\omega$ by a subset of a set of $\mathcal{F}_\omega$ of measure 0, and we define $\mu(\Lambda_1)$ as $\mu(\Lambda_0)$. Let $\mathcal{F}^*$ be the following space: a point of $\mathcal{F}^*$ is a class of measurable ω-sets, any two of which differ at most by a set of measure 0. We metrize $\mathcal{F}^*$ as follows: if Λ_1^*, Λ_2^* are points of $\mathcal{F}^*$, and if Λ_i is a set in the class Λ_i^*, the distance between Λ_1^*, Λ_2^* is defined as arc tan $\mu(\Lambda_1 + \Lambda_2 - \Lambda_1 \cdot \Lambda_2)$,[5] or $\frac{1}{2}\pi$ if $\mu(\Lambda_1 + \Lambda_2 - \Lambda_1 \cdot \Lambda_2)$

Received May 7, 1938.

[1] A field of sets is a collection of sets including $E_1 + E_2$, $E_1 - E_1 \cdot E_2$ if it includes E_1, E_2. A Borel field of sets is a field of sets including $\sum_1^\infty E_n$ if it includes $E_1, E_2, \cdots$.

[2] The Borel field of sets determined by a given collection of sets is the smallest Borel field of sets containing the given collection.

[3] The collection of all sets $\tilde{\Lambda}$ having this property is readily seen to be a Borel field. This Borel field certainly includes every set $E \times \Lambda$ as defined above, so the field is precisely $\mathcal{F}_t \times \mathcal{F}_\omega$.

[4] If $\mu(\Lambda)$ is not always finite-valued, we assume that Ω can be expressed as a denumerably infinite sum of sets in $\mathcal{F}_\omega$ on each of which $\mu(\Lambda)$ is finite-valued.

[5] If $\mu(\Omega) < +\infty$, we can define the distance between Λ_1^*, Λ_2^* as $\mu(\Lambda_1 + \Lambda_2 - \Lambda_1 \cdot \Lambda_2)$ and obtain the same $\mathcal{F}^*$-topology, but this definition is not possible in the general case, if the distance function is to be finite-valued.

$= +\infty$. Any collection of measurable sets will be called *separable* if the corresponding $\mathfrak{F}^*$-point set is a separable point set (i.e., if the corresponding $\mathfrak{F}^*$-point set has a denumerable subset dense on it). A strictly separable Borel field is separable if $\mu(\Omega) < +\infty$; a separable Borel field need not be strictly separable.

We shall also suppose that there is a non-negative completely additive function mE of sets $E \in \mathfrak{F}_t$ and that if $\mu(T) = +\infty$, T is the sum of a denumerable sequence of sets of finite measure. We extend the definition of mE to measurable t-sets in the usual way—as $\mu(\Lambda)$ was extended.

If $\tilde{\Lambda}$ is a (t, ω)-set of the form $E \times \Lambda$, $E \in \mathfrak{F}_t$, $\Lambda \in \mathfrak{F}_\omega$, we set $\tilde{\mu}(\tilde{\Lambda}) = mE \cdot \mu(\Lambda)$. It is well known[6] that this definition determines a completely additive non-negative set function on the sets of $\mathfrak{F}_t \times \mathfrak{F}_\omega$, and this set function is extended to the measurable (t, ω)-sets in the usual way.

We shall use repeatedly the fact that whenever a Borel field of sets is extended to a field of measurable sets, as described above, if f_1 is a numerically-valued measurable function, there is a function f_0, equal to f_1 almost everywhere, and measurable with respect to the given Borel field. This fact is still true if f_1 is no longer supposed numerically-valued, but if it is supposed to take on values in a separable Hausdorff space—measurability of such a function meaning that open f-sets correspond to measurable sets of the argument space.

THEOREM 1. *Let $\mu(\Omega) < +\infty$ and let $f(t, \omega)$ be a numerically-valued function which is ω-measurable for almost all values of t. There is a (t, ω)-measurable function $f_0(t, \omega)$ such that, for each value of t, $f(t, \omega) = f_0(t, \omega)$ almost everywhere on Ω if and only if*

(i) *there is a t-set S_0 of measure 0 and a separable field $\mathfrak{F}_\omega(f)$ of measurable ω-sets such that $f(t, \omega)$, for t fixed, $t \in' S_0$,*[7] *is measurable with respect to the field $\mathfrak{F}_\omega(f)$, and*

(ii) *there is a function $\varphi(y)$, a monotone increasing function (in the strict sense) of the real variable y, such that if Λ is any measurable ω-set, $\varphi[f(t, \omega)]$ is ω-integrable on Λ for almost all t, and*

$$\int_\Lambda \varphi[f(t, \omega)]\, d\omega$$

is a measurable function of t.[8]

[6] Cf. S. Saks, *Theory of the Integral*, New York, 1937, pp. 82–88.

[7] The symbol $\in'$ means "is not an element of".

[8] If (i) is true, and if there is such a function $\varphi(y)$, then if $\varphi_1(y)$ is any Baire function of y such that, if Λ is any measurable ω-set of finite measure, $\varphi_1[f(t, \omega)]$ is integrable on Λ for almost all t, $\int_\Lambda \varphi_1[f(t, \omega)]\, d\omega$ will be a measurable function of t, since

$$\int_\Lambda \varphi_1[f(t, \omega)]\, d\omega = \int_\Lambda \varphi_1[f_0(t, \omega)]\, d\omega,$$

and the latter is a measurable function of t, by Fubini's theorem. A special case of Theorem 1 was proved by Bochner and von Neumann, Annals of Mathematics, vol. 36(1935), pp. 263–265.

If a function $f_0(t, \omega)$ exists, with the properties described, there is a function $f_1(t, \omega)$ which is measurable with respect to the field $\mathcal{F}_t \times \mathcal{F}_\omega$ and equal to $f_0(t, \omega)$ for almost all (t, ω). It is readily seen that there are separable subfields $\mathcal{F}_t'$ of $\mathcal{F}_t$, $\mathcal{F}_\omega'$ of $\mathcal{F}_\omega$ such that $f_1(t, \omega)$ is measurable with respect to the field $\mathcal{F}_t' \times \mathcal{F}_\omega'$. From Fubini's theorem we know that there is a t-set S_0 of measure 0 such that, if $t \epsilon' S_0$, $f_1(t, \omega) = f_0(t, \omega)$ almost everywhere on Ω. Then if t is fixed and if $t \epsilon' S_0$, $f(t, \omega) = f_1(t, \omega)$ almost everywhere on Ω, and $f(t, \omega)$ is measurable with respect to the field $\mathcal{F}_\omega'$ increased by all sets differing from a set in $\mathcal{F}_\omega'$ by at most a set of measure 0. This extended field can be taken as the field $\mathcal{F}_\omega(f)$ of (i). In (ii) let $\varphi(y) = \arctan y$ or any other bounded monotone increasing function of y. Then if Λ is a measurable ω-set,

$$\int_\Lambda \varphi[f_0(t, \omega)]\, d\omega \equiv \int_\Lambda \varphi[f(t, \omega)]\, d\omega$$

is a measurable function of t, by Fubini's theorem.

Conversely, suppose that (i) and (ii) are true. To prove the existence of the function $f_0(t, \omega)$ as described, it is sufficient to assume that $\varphi(y)$ is bounded, for if this were not so, we could replace $\varphi(y)$ by $\arctan \varphi(y)$, and (ii) would still be true. Replacing now $f(t, \omega)$ by $\varphi[f(t, \omega)]$, we see that it is sufficient to prove that if there is a t-set S_0 as described in (i), if $f(t, \omega)$ is bounded, and if $\int_\Lambda f(t, \omega) d\omega$ is a measurable function of t for every measurable ω-set Λ, there is a function $f_0(t, \omega)$ as described in the theorem. Since $\mathcal{F}_\omega(f)$ is separable, the complex-valued functions, measurable with respect to $\mathcal{F}_\omega(f)$, whose moduli squared are integrable, determine a finite-dimensional unitary space or a Hilbert space, making the usual inner product definition.[9] Let $\{g_n(\omega)\}$ be a complete normal orthogonal set of functions. If $t \epsilon' S_0$, the series

$$\sum_1^\infty a_n(t) g_n(\omega), \qquad a_n(t) = \int_\Omega f(t, \omega)\overline{g_n(\omega)}\, d\omega,$$

converges in measure to $f(t, \omega)$ on Ω.[10] Now by hypothesis, if $g(\omega)$ is the characteristic function of a measurable ω-set,

$$\int_\Omega f(t, \omega)\overline{g(\omega)}\, d\omega$$

is a measurable function of t. The integral will therefore be measurable (if we use the familiar approximation of a measurable function by linear combinations of characteristic functions) whenever $g(\omega)$ is measurable and integrable, so that

[9] Cf., for example, M. H. Stone, *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publications, vol. 15, 1932, Chapter I, §5.

[10] Convergence in measure is discussed by E. W. Hobson, *Theory of Functions of a Real Variable*, vol. 2, 2d edition, pp. 240–245. The convergence in measure here follows at once from the convergence in the mean (Riesz-Fischer theorem).

each $a_n(t)$ is a measurable function of t. The (t, ω)-function $a_n(t)g_n(\omega)$ is thus the product of a measurable function of t and a measurable function of ω and is therefore a measurable function of (t, ω). The partial sums

$$\sum_{n=1}^{N} a_n(t)g_n(\omega) \qquad (N = 1, 2, \cdots)$$

are then measurable functions of (t, ω). If S is a measurable t-set of finite measure, it follows readily[11] that the partial sums converge in measure on $S \times \Omega$ and therefore that some subsequence of partial sums converges almost everywhere on $S \times \Omega$. Since S was an arbitrary measurable t-set of finite measure, and since T is the sum of denumerably many such sets, a familiar application of the diagonal procedure provides a subsequence of partial sums converging almost everywhere on $T \times \Omega$ to some measurable function of (t, ω), $f_1(t, \omega)$. There is a t-set S_1 of measure 0 such that if $t \epsilon' S_1$, this subsequence of partial sums converges almost everywhere on Ω to $f_1(t, \omega)$. Then if $t \epsilon' S_0 + S_1$, $f_1(t, \omega) = f(t, \omega)$ almost everywhere on Ω. We can obtain the function $f_0(t, \omega)$ of the theorem by changing $f_1(t, \omega)$, for $t \epsilon S_0 + S_1$, to be equal to $f(t, \omega)$.

We shall need the following lemma.

LEMMA 1. *Let t-space be the space of real numbers, and let t-measure be Lebesgue measure. Let $\Phi(t)$ be a function defined for almost all t, and taking on values in a metric space. Then if $\Phi(t)$ is approximately continuous for almost all t,*[12]

(i) *$\Phi(t)$ is measurable;*

(ii) *if G is any finite t-interval, there is a measurable subset P of G of measure arbitrarily near that of G at the points of which $\Phi(t)$ is continuous relative to P;*

(iii) *there is a t-set S_0 of measure 0 such that the values of $\Phi(t)$ for $t \epsilon' S_0$ form a separable Φ-set.*

Conversely, if (i) *and* (iii) *are true, $\Phi(t)$ is approximately continuous for almost all t.*

Suppose that $\Phi(t)$ is approximately continuous for almost all t. Then it has been shown by Stepanoff[13] and Kamke[14] that $\Phi(t)$ is measurable. To prove (ii)—that Lusin's theorem holds—let G be any open t-interval, and let N be a positive integer. If $\Phi(t)$ is approximately continuous at t_0, there is an open interval $I(t_0)$ such that $t_0 \epsilon I(t_0)$ and for any subinterval I of $I(t_0)$ containing t_0, there is a measurable t-set $E(t_0) \subseteq I(t_0)$ which satisfies $mI \cdot E(t_0) \geqq (1 - 2^{-N})^{\frac{1}{2}} mI$ and on which the oscillation of $\Phi(t)$ is at most N^{-1}. According

[11] Cf. J. L. Doob, Transactions of the American Mathematical Society, vol. 42(1937), pp. 114–115.

[12] We take this to mean that for almost all values of t, $\Phi(t + h) \rightarrow \Phi(t)$ when $h \rightarrow 0$ on some set (depending on t) having metric density 1 at $h = 0$.

[13] Recueil Mathématique de Moscou, vol. 31(1924), pp. 487–489.

[14] Fundamenta Mathematicae, vol. 10(1927), pp. 431–433. These proofs assume that $\Phi(t)$ is a numerically-valued real function, but are applicable to the present case.

to Vitali's theorem,[15] there is a finite number of such intervals $I: I_1^N, I_2^N, \cdots$ corresponding to points $t_1^{(N)}, t_2^{(N)}, \cdots$, so that

$$mI_j^N \cdot E(t_j^{(N)}) \geqq (1 - 2^{-N})^{\frac{1}{2}} mI_j^N,$$

where the intervals $I_1^N, I_2^N, \cdots$ are disjunct and

$$\sum_j mI_j^N \geqq (1 - 2^{-N})^{\frac{1}{2}} mG.$$

Then if

$$E_N = \sum_j E(t_j^{(N)}) \cdot I_j^N,$$

it follows that

$$mE_N = \sum_j mE(t_j^{(N)}) \cdot I_j^N \geqq (1 - 2^{-N})^{\frac{1}{2}} \sum_j mI_j^N \geqq (1 - 2^{-N}) mG$$

and the oscillation of $\Phi(t)$ at each point of E_N, relative to E_N, is at most N^{-1}. Let E^N be the common part of the sets $E_N, E_{N+1}, \cdots$:

$$E^N = \prod_N^\infty E_n.$$

Then[16]

$$mE^N = mG - mG \cdot CE^N = mG - m\left(G \cdot \sum_N^\infty CE_n\right) \geqq (1 - 2^{-N+1}) mG$$

and $\Phi(t)$ has oscillation 0 relative to E^N, at every point of E^N; i.e., $\Phi(t)$ is continuous on E^N relative to E^N, and E^N has measure arbitrarily close to mG, since N is arbitrary, as was to be proved. To prove (iii) let P_n be a measurable t-set, in $|t| < n$, of measure at least $2n - n^{-1}$, such that $\Phi(t)$ is continuous everywhere on P_n relative to P_n. Then if $t_1, t_2, \cdots$ is a sequence of values of t, including for every value of n a subsequence lying in P_n and dense on P_n, the values $\{\Phi(t_j)\}$ form a Φ-set everywhere dense on the values of $\Phi(t)$ assumed in $\sum_1^\infty P_n$, and (iii) is true with S_0 the complement of $\sum_1^\infty P_n$.

Conversely, suppose that (i) and (iii) are true. Then a slight modification (due to the fact that $\Phi(t)$ is not numerically-valued) of L. W. Cohen's proof[17] of Lusin's theorem shows that (ii) is true, and this implies that $\Phi(t)$ is almost everywhere approximately continuous, if we use Lebesgue's theorem on metric density.

In the following theorem, we develop Theorem 1 further, considering only functions $f(t, \omega)$ which are characteristic functions of point sets, to simplify the presentation.

THEOREM 2. *Let $\tilde{\Lambda}$ be a (t, ω)-set, and let Λ_{t_0} be the cross-section of $\tilde{\Lambda}$ at $t = t_0$.*

[15] S. Saks, *Theory of the Integral*, New York, 1937, pp. 109–111.

[16] The prefix C will be used to denote the operation of taking complements.

[17] *Fundamenta Mathematicae*, vol. 9(1927), pp. 122–123.

Suppose that Λ_t is a measurable ω-set for almost all t. Let $\Phi^{\mathrm{M}}(t)$ be the function of t taking on values in the metric space $\mathfrak{F}^*$ defined above, for which $\Phi^{\mathrm{M}}(t_0)$ is the point of $\mathfrak{F}^*$ corresponding to $\Lambda_{t_0}\cdot\mathrm{M}$, where, throughout this theorem, M is a measurable ω-set of finite measure if $\mu(\Omega) = \infty$, and otherwise $\mathrm{M} = \Omega$. The following statements are equivalent.

(i) *For each M, $\Phi^{\mathrm{M}}(t)$ is a measurable function of t, and there is a t-set S_0 (independent of M) of measure 0, such that the sets $\Lambda_t\cdot\mathrm{M}$, $t \epsilon' S_0$, form a separable collection.*

(ii) *$\tilde{\Lambda}$ can be changed into a measurable (t, ω)-set by changing Λ_t for each t by an ω-set of measure 0.*

(iii) *$\mu(\Lambda_t\cdot\Lambda)$ is a measurable function of t, for every measurable ω-set Λ, and there is a t-set S_0, as in* (i).

(iv) *If t-space is the space of real numbers, and if t-measure is Lebesgue measure, except for a t-set S_0 (independent of M) of measure 0,*

$$\mu(\mathrm{M}\cdot\Lambda_{t+h}\cdot C\Lambda_t) + \mu(\mathrm{M}\cdot\Lambda_t\cdot C\Lambda_{t+h}) = \mu((\Lambda_t + \Lambda_{t+h} - \Lambda_t\cdot\Lambda_{t+h})\mathrm{M}) \to 0,$$

for each M, when $h \to 0$ on a set (which may depend on t but not on M) having metric density 1 at $h = 0$.[18]

It will be sufficient to show the equivalence of (i) and (iii), of (ii) and (iii), and of (i) and (iv). We discuss only the case $\mu(\Omega) < \infty$ in which $\mathrm{M} = \Omega$ throughout the above statements. The general case follows readily, if we use the fact that, if $\mu(\Omega) = \infty$, we have supposed that Ω is the sum of denumerably many measurable ω-sets of finite measure.

Proof that (i) $\rightleftarrows$ (iii). Suppose that (i) is true. Let Λ_t^* be the $\mathfrak{F}^*$-point corresponding to the ω-set Λ_t. There is then a sequence of t-values $t_1, t_2, \cdots, t_n \epsilon' S_0$, such that the $\mathfrak{F}^*$-points $\Lambda_{t_1}^*, \Lambda_{t_2}^*, \cdots$ are dense on the $\mathfrak{F}^*$-point set $\{\Lambda_t^*\}$, $t \epsilon' S_0$. Let ν be a positive integer. Define $\Phi_\nu(t)$ as the function taking on values in $\mathfrak{F}^*$ which for $t = \tau \epsilon' S_0$ is equal to that $\Lambda_{t_j}^*$ with smallest j which has distance $\leq \nu^{-1}$ from Λ_τ^*. The function $\Phi_\nu(t)$ is defined for almost all values of t, takes on at most denumerably many values, each on a measurable t-set (since $\Phi(t) = \Phi^{\Omega}(t)$ is measurable by hypothesis), and is therefore measurable. If $t \epsilon' S_0$, let Λ_t^ν be an ω-set corresponding to $\Phi_\nu(t)$. Then if Λ is a measurable ω-set, $\mu(\Lambda_t^\nu\cdot\Lambda)$ takes on at most denumerably many values, each on a measurable

[18] Statement (iv) was suggested to the author by a theorem in probability on the properties of stochastic processes depending on a continuous parameter, which was stated in a letter from A. Kolmogoroff (the terminology used was developed by the present author in the Transactions of the American Mathematical Society, vol. 42(1937), pp. 107–140): *A necessary and sufficient condition that there is a measurable stochastic process with a given P^*-measure is that, if ϵ is any positive number,*

$$\lim_{h\to 0} P^*\{|x(t+h) - x(t)| > \epsilon\} = 0$$

for almost all t, when $h \to 0$ on a set (which may depend on t) having metric density 1 at $h = 0$. Kolmogoroff's theorem can be proved from the equivalence of (ii) and (iv) of the present theorem.

t-set, so $\mu(\Lambda_t^\nu\cdot\Lambda)$ is a measurable function of t. When $\nu\to\infty$, $\mu(\Lambda_t^\nu\cdot\Lambda)\to\mu(\Lambda_t\cdot\Lambda)$, so $\mu(\Lambda_t\cdot\Lambda)$ is also a measurable function of t, and (iii) is true. Conversely, suppose that (iii) is true, so that $\mu(\Lambda_t\cdot\Lambda)$ is a measurable function of t for every measurable ω-set Λ, and that there is a t-set S_0, as described in (i). To show that $\Phi(t) = \Phi^\Omega(t)$ is measurable, it suffices to show that the t-set corresponding (through $\Phi(t)$) to any open $\mathfrak{F}^*$-set is measurable. Since (neglecting t-values in S_0) the values assumed by $\Phi(t)$ form a separable set, it is sufficient to prove that the t-set corresponding to any $\mathfrak{F}^*$-sphere is measurable. This means, if ρ is the radius of the sphere, that it must be shown that, if Λ is a measurable ω-set (corresponding to the center of the sphere), the t-set defined by the inequality

$$\mu(\Lambda + \Lambda_t - \Lambda\cdot\Lambda_t) = \mu(\Lambda_t\cdot C\Lambda) + \mu(\Lambda\cdot C\Lambda_t) < \tan\rho$$

is measurable. Now $\mu(\Lambda_t\cdot\Lambda)$ is a measurable function of t, by hypothesis, and $\mu(\Lambda\cdot C\Lambda_t) = \mu(\Lambda) - \mu(\Lambda_t\cdot\Lambda)$ is also a measurable function of t, so that $\mu(\Lambda + \Lambda_t - \Lambda\cdot\Lambda_t)$ is a measurable function of t. The inequality $\mu(\Lambda + \Lambda_t - \Lambda\cdot\Lambda_t) < \tan\rho$ therefore defines a measurable t-set, as was to be proved.

Proof that (ii) $\rightleftarrows$ (iii). This is essentially Theorem 1 for a function $f(t,\omega)$ taking on only the values 1 and 0, except that we have allowed $\mu(\Omega)$ to be $+\infty$. This extension presents no difficulty.

Proof that (i) $\rightleftarrows$ (iv). We suppose now that t-space is the space of real numbers, and that t-measure is Lebesgue measure. Condition (iv) (if $\mu(\Omega)<\infty$) is simply the condition that $\Phi(t)$ be almost everywhere approximately continuous. If (i) is true, $\Phi(t)$ is almost everywhere approximately continuous, according to Lemma 1, so (iv) is true. Conversely if (iv) is true, so that $\Phi(t)$ is almost everywhere approximately continuous, (i) is true, also by Lemma 1.

If $\mu(\Lambda_t\cdot\Lambda)$ is continuous at t_0 for $\Lambda = C\Lambda_{t_0}$,

$$\lim_{h\to 0}\mu(\Lambda_{t_0+h}\cdot C\Lambda_{t_0}) = \mu(\Lambda_{t_0}\cdot C\Lambda_{t_0}) = 0,$$

and if $\mu(\Lambda_t\cdot\Lambda)$ is continuous at t_0 for $\Lambda = \Lambda_{t_0}$,

$$\lim_{h\to 0}\mu(\Lambda_{t_0}\cdot C\Lambda_{t_0+h}) = \lim_{h\to 0}[\mu(\Lambda_{t_0}) - \mu(\Lambda_{t_0+h}\cdot\Lambda_{t_0})] = 0$$

if $\mu(\Lambda_{t_0})<\infty$. This suggests an alternative statement of (iv) which will be useful in the applications:

(iv′) *If t-space is the space of real numbers and if t-measure is Lebesgue measure, $\mu(\Lambda_t\cdot M)$, for each M, is an approximately continuous function of t, except possibly for a t-set of measure* 0, *independent of* M.

It is clear that (iv) and (iv′) are equivalent if $\mu(\Omega)<\infty$. The proof for $\mu(\Omega)=\infty$ is then readily obtained if we use the fact that in this case Ω is the sum of denumerably many sets of finite measure.

LEMMA 2. *Let t-space be the space of real numbers, and let t-measure be Lebesgue measure. Let $f(t,\omega)$ be a measurable function of (t,ω) taking on values in a separable Hausdorff space, and defined for $t\neq t_0$, and suppose that $f(t,\omega)$ is measurable*

in ω for each fixed value of t. Suppose that there is a function $f(\omega)$ such that, whenever $\{t_n\}$ is a sequence converging to t_0, $\lim_{n\to\infty} f(t_n, \omega) = f(\omega)$ almost everywhere on Ω. Then if $\{s_n\}$ is any denumerably infinite t-set having t_0 as a limit point, $\lim_{t\to t_0} f(t, \omega) = f(\omega)$ almost everywhere on Ω, when $t \to t_0$ on the set $\{s_n\}$.

It is evidently sufficient to consider only the case $\mu(\Omega) < +\infty$.

Since the f-space is supposed separable, there is a sequence of open f-sets: $F_1, F_2, \cdots$, such that any open f-set is the sum of the F_j-sets contained in it. Consider the ω-set $\Lambda_{n,N}$, consisting of the points for which $f(\omega) \in F_N$, but for which $f(s_j, \omega) \in' F_N$ for some point s_j satisfying the inequality $0 < |s_j - t_0| < n^{-1}$:

$$\Lambda_{n,N} = \sum_{0<|s_j-t_0|<1/n} \{f(\omega) \in F_N, f(s_j, \omega) \in' F_N\}.$$

The set $\Lambda_{n,N}$ is measurable and does not increase when n increases. We shall show that

$$\lim_{n\to\infty} \mu(\Lambda_{n,N}) = 0 \qquad (N = 1, 2, \cdots).$$

Unless this is true, there is an integer N and a positive number η such that

$$\mu(\Lambda_{n,N}) \geqq \eta \qquad (n = 1, 2, \cdots).$$

Let $s_1^{(n)}, s_2^{(n)}, \cdots$ be the values of s_n such that $0 < |s_n - t_0| < n^{-1}$. Then

$$\Lambda_{n,N} = \sum_j \{f(\omega) \in F_N, f(s_j^{(n)}, \omega) \in' F_N\}.$$

There is an integer ν_n so large that

$$\mu\left(\Lambda_{n,N} - \sum_{j=1}^{\nu_n} \{f(\omega) \in F_N, f(s_j^{(n)}, \omega) \in' F_N\}\right) < \eta 2^{-n-1}.$$

Let $\Lambda'_{n,N}$ be the set

$$\sum_{j=1}^{\nu_n} \{f(\omega) \in F_N, f(s_j^{(n)}, \omega) \in' F_N\}.$$

Then the set $\Lambda'_N = \prod_{n=1}^{\infty} \Lambda'_{n,N}$ is of positive measure:

$$\mu(\Lambda'_N) \geqq \mu\left[\prod_{n=1}^{\infty} \Lambda_{n,N} - \prod_{n=1}^{\infty} \Lambda_{n,N} \sum_{n=1}^{\infty} (\Lambda_{n,N} - \Lambda'_{n,N})\right] \geqq \eta - \eta \sum_{1}^{\infty} 2^{-n-1} = \tfrac{1}{2}\eta.$$

Let $t_1, t_2, \cdots$ be the sequence $s_1^{(1)}, \cdots, s_{\nu_1}^{(1)}, s_1^{(2)}, \cdots, s_{\nu_2}^{(2)}, \cdots$. Then $t_n \to t_0$, and if $\omega \in \Lambda'_N$, $f(\omega) \in F_N$, but $f(t_j, \omega) \in' F_N$ for infinitely many values of j, so that, when $j \to \infty$, $f(t_j, \omega)$ does not converge to $f(\omega)$ on Λ'_N. This contradicts the hypotheses, so the assumption $\mu(\Lambda_{n,N}) \nrightarrow 0$ is false. Now for fixed ω the statement

$$\lim_{t\to t_0} f(t, \omega) = f(\omega) \qquad (t \in \{s_n\})$$

can be restated: if F_N is any F_j-set containing $f(\omega)$, $f(t, \omega) \epsilon F_N$ for sufficiently small values of $t - t_0$ ($t \epsilon \{s_n\}$). Then if $\{n_N\}$ is any sequence of positive integers, the ω-set where it is not true that $\lim_{t \to t_0} f(t, \omega) = f(\omega)$ ($t \epsilon \{s_n\}$) is included in the set $\sum_{N=1}^{\infty} \Lambda_{n_N, N}$. If $\epsilon > 0$, since $\lim_{n \to \infty} \mu(\Lambda_{n,N}) = 0$, we can choose an integer n_N, for each value of N, so large that $\mu(\Lambda_{n_N, N}) < \epsilon 2^{-N}$. Then

$$\mu\left(\sum_{N=1}^{\infty} \Lambda_{n_N, N}\right) < \epsilon,$$

so that the ω-set of non-convergence has measure at most ϵ, and therefore measure 0, as was to be proved.

THEOREM 3. *Suppose that t-space is the space of real numbers and that t-measure is Lebesgue measure. Let $f(t, \omega)$ be a measurable function of (t, ω), taking on values in a separable Hausdorff space. Suppose that there is a t-set S_0 of measure 0 such that, whenever $t_n \to t \epsilon' S_0$, $f(t_n, \omega) \to f(t, \omega)$ almost everywhere on Ω. Then, neglecting an ω-set of measure 0, $f(t, \omega)$, for fixed ω, is a continuous function of t relative to a t-set (depending on ω) whose complement is of measure 0.*[19]

Let $F_1, F_2, \cdots$ be a sequence of open sets of f-space such that any open set of this space is the sum of the F_j-sets contained in it. Define $\varphi_n^N(t, \omega)$ as follows: let ω be fixed and suppose that $k2^{-n} \leqq t < (k + 1)2^{-n}$; then $\varphi_n^N(t, \omega) = 1$ if $f(t, \omega) \epsilon F_N$ for every rational value of t in the interval $(k - 1)2^{-n} < t < (k + 1)2^{-n}$, and otherwise, $\varphi_n^N(t, \omega) = 0$. Evidently $\varphi_n^N(t, \omega)$ is a measurable function of (t, ω). Moreover, $\varphi_1^N \leqq \varphi_2^N \leqq \cdots$, so that $\lim_{n \to \infty} \varphi_n^N = \varphi^N$ exists for all (t, ω) and is (t, ω)-measurable. By Lemma 2, if $t' \epsilon' S_0$, and if $r \to t'$ (r rational), $f(r, \omega) \to f(t', \omega)$ almost everywhere on Ω, so that if $f(t', \omega) \epsilon F_N$, $f(r, \omega) \epsilon F_N$ for r sufficiently near t' (neglecting an ω-set of measure 0). Then if $f(t', \omega) \epsilon F_N$, $\varphi_n^N(t', \omega) = 1$ for n sufficiently large: $\varphi^N(t', \omega) = 1$ almost everywhere on the set Λ_N where $f(t', \omega) \epsilon F_N$. In other words if $f^N(t, \omega)$ is the characteristic function of Λ_N, $\varphi^N(t, \omega) = f^N(t, \omega)$ almost everywhere on Λ_N. Then neglecting an ω-set $\Lambda(F_N)$ of measure 0, if ω is fixed, in Λ_N, $\varphi^N(t, \omega) = f^N(t, \omega)$ for almost all t: $t \epsilon' S(\omega, F_N)$. Now from the definition of φ^N, if $\varphi^N(t', \omega) = 1$, $\varphi^N(t, \omega) = 1$ for t near t' (ω fixed), so if ω is fixed, not in $\Lambda(F_N)$, and if $t \epsilon' S(\omega, F_N)$, whenever $f(t', \omega) \epsilon F_N$, $f(t, \omega) \epsilon F_N$ for t near t'. The theorem states that (neglecting an ω-set Λ of measure 0) if $t_0 \epsilon' S(\omega)$, where $S(\omega)$ is of measure 0, $f(t, \omega)$ is continuous at t_0 relative to the complement of $S(\omega)$. This can be verified at once, if we set

$$S(\omega) = \sum_{N=1}^{\infty} S(\omega, F_N), \qquad \Lambda = \sum_{N=1}^{\infty} \Lambda(F_N).$$

[19] This theorem can be interpreted to give an important result in the theory of stochastic processes depending on a continuous parameter. Cf. J. L. Doob, Transactions of the American Mathematical Society, vol. 42(1937), p. 118.

Families of set transformations

Throughout the rest of this paper, t-space will be supposed to be the space of real numbers, $\mathfrak{F}_t$ the field of Borel sets, and t-measure Lebesgue measure. The families of transformations to be discussed are of two types.

I. In the first type we suppose that there is a given space Ω and a Borel field $\mathfrak{F}_\omega$ of ω-sets, as discussed above. We suppose that set transformations $T_t\Lambda$ are defined, satisfying the following conditions:

(α) If $\Lambda \in \mathfrak{F}_\omega$, $T_t\Lambda$ are defined for almost all values of t and are sets of $\mathfrak{F}_\omega$. If $T_s\Lambda$ and $T_t(T_s\Lambda)$ are defined, then $T_{t+s}\Lambda$ is also defined and is $T_t(T_s\Lambda)$.

(β) If $\Lambda_1, \Lambda_2, \cdots$ is any sequence of sets of $\mathfrak{F}_\omega$, and if $T_t\Lambda_n$ is defined for all n, then $T_t\left(\sum_1^\infty \Lambda_n\right)$ and $T_t\left(\prod_1^\infty \Lambda_n\right)$ are defined, and

$$T_t\left(\sum_1^\infty \Lambda_n\right) = \sum_1^\infty T_t\Lambda_n, \qquad T_t\left(\prod_1^\infty \Lambda_n\right) = \prod_1^\infty T_t\Lambda_n.$$

If there is a point transformation which induces the set transformation T_t, the first half of (β) is automatically satisfied.

The second type of family to be considered is probably more important.

II. In the second type we suppose not only the space Ω and field $\mathfrak{F}_\omega$, but a measure $\mu(\Lambda)$, as discussed above, giving rise to a field of ω-measurable sets. Let Λ be any measurable ω-set. Then we suppose that $T_t\Lambda$ is defined for almost all t and is a measurable ω-set, and that if $\mu(\Lambda) = 0$, then $\mu(T_t\Lambda) \equiv 0$. We no longer suppose T_t to be uniquely defined, but only up to an ω-set of measure 0. We suppose that if $T_t\Lambda_1$ is defined, and if Λ_2 differs from Λ_1 by at most an ω-set of measure 0, then $T_t\Lambda_2$ is also defined and is $T_t\Lambda_1$, aside from the ambiguity just mentioned. Hypotheses (α) and (β) are made as in I, with the obvious changes forced by the multiple-valued nature of the transformations.

It follows from (α) and (β) that if $\mathfrak{F}'_\omega$ is any strictly separable subfield of $\mathfrak{F}_\omega$, there is a t-set S_0 of Lebesgue measure 0 such that $T_t\Lambda$ is defined for Λ in $\mathfrak{F}'_\omega$ and t not in S_0. In Case II this conclusion also holds with $\mathfrak{F}'_\omega$ any separable collection of measurable ω-sets.

Before going into a discussion of families of transformations, we shall give several examples.

(a) Let the space Ω be the perimeter of a circle, and let $\mathfrak{F}_\omega$ be the field of all point sets on this perimeter. We define T_t as the rotation through t radians. The example comes under Case I. The field $\mathfrak{F}_\omega$ is not strictly separable.

(b) Let the space Ω be the (x, y)-plane and let $\mathfrak{F}_\omega$ be the field of Borel planar sets. We define T_t as the translation $x' = x + t$. Let $f(x, y)$ be a non-negative Baire function, and define $\mu(\Lambda)$, if Λ is a Borel planar set, by

$$\int\int_\Lambda f(x, y)\, dx\, dy.$$

Then the family comes under Case II, if $f(x, y)$ never vanishes.

(c) Let the space Ω be a subset of the perimeter of a circle of radius $(2\pi)^{-1}$, of exterior Lebesgue measure 1. Let $\mathfrak{F}_\omega$ be the field of intersections of Ω with the Borel sets of the perimeter. If E is a Borel set of the perimeter, the ω-measure of $E \cdot \Omega$ is defined as the Borel measure of E. It can be shown that this defines ω-measure uniquely.[20] Let T_t be the set transformation taking the intersection $E \cdot \Omega$ of the Borel set E of the perimeter with Ω into $E_t \cdot \Omega$, where E_t is E rotated through t radians. Evidently if T_t is generated by a point transformation, the point transformation must be a rotation through t radians, and this presupposes that Ω is invariant under such a rotation, a supposition not necessarily fulfilled.

(d) Let the space Ω be the space of all real-valued functions $x(t)$, $-\infty < t < +\infty$. Let $F(\lambda)$ be any monotone non-decreasing function satisfying the following conditions:

$$\lim_{\lambda \to -\infty} F(\lambda) = 0, \qquad \lim_{\lambda \to +\infty} F(\lambda) = 1, \qquad \lim_{\epsilon \to 0} F(\lambda + \epsilon^2) = F(\lambda),$$

and suppose that $F(\lambda)$ takes on at least one value not 0 or 1. Consider the ω-set defined by the conditions

$$x(t_j) \leqq \lambda_j \qquad (j = 1, \cdots, n),$$

where n is any natural number, $t_1, \cdots, t_n$ any n distinct values of t, $\lambda_1, \cdots, \lambda_n$ any n real numbers. To this set is assigned the measure $\prod_1^n F(\lambda_j)$. If $\mathfrak{F}_\omega$ is the Borel field determined by all such sets, ω-measure is completely determined.[21] In the language of probability, the values of $x(t)$ at the different values of t are independent of each other. The transformation T_τ is the transformation taking $\omega: x(t)$ into $T_\tau \omega: x(t + \tau)$. The measurable ω-sets do not form a separable collection. For if λ_0 is so chosen that $0 < F(\lambda_0) < 1$, the condition $x(t) \leqq \lambda_0$ defines a measurable ω-set $\Lambda_t = T_{-t}\Lambda_0$, and the $\mathfrak{F}^*$-distance between any two of these sets $\Lambda_{t_1}, \Lambda_{t_2}$ is

$$d = \text{arc tan } \{2F(\lambda_0)[1 - F(\lambda_0)]\} > 0,$$

independent of t_1, t_2.

Let $T \times \Omega$ and (t, ω)-measure be defined as in the introduction. Suppose that a family of set transformations T_t is given, as defined above in Case I. Let Λ be a set of $\mathfrak{F}_\omega$ and consider a (t, ω)-set $\tilde{\Lambda}$ which meets $t = t_0$ in $T_{t_0}\Lambda$, if $T_{t_0}\Lambda$ is defined, and is otherwise arbitrary. If $\tilde{\Lambda}$ is a measurable (t, ω)-set for every Λ in $\mathfrak{F}_\omega$, the family of transformations will be called *measurable*.[22] In

[20] Cf. J. L. Doob, Transactions of the American Mathematical Society, vol. 42(1937), pp. 109–110.

[21] A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik, vol. 2, No. 3, pp. 24–30.

[22] This definition generalizes one given by von Neumann, Annals of Mathematics, vol. 33(1932), p. 589, who discusses families of measure-preserving point transformations. It should be noted that the definition depends on the particular ω-measure defined. If $\tilde{\Lambda}$ is a set of $\mathfrak{F}_t \times \mathfrak{F}_\omega$, for each Λ in $\mathfrak{F}_\omega$, however, the family of transformations will be measurable whatever the definition of ω-measure.

Case II we have allowed the definition of T_t to be ambiguous. In defining measurability we shall require that one possible (t, ω)-set $\tilde{\Lambda}$ corresponding to the given set Λ be a measurable (t, ω)-set, if Λ is any ω-measurable set.

It is easily shown that, if a family of set transformations is being considered, under Case I, and if the set transformations are generated by point transformations (which we shall denote by $T_t\omega$), the family of transformations is measurable if and only if, whenever $f(\omega)$ is a real-valued function of ω measurable with respect to the field $\mathcal{F}_\omega$, $f(T_t\omega)$ is a (t, ω)-measurable function.

THEOREM 4. *Let* $\{T_t\}$ *be a measurable family of set transformations, Case* I *or* II. *In the following,* M *will always be a measurable ω-set of finite measure, and* $\mathrm{M} \equiv \Omega$ *in* (i), (iii), (iv), *if* $\mu(\Omega) < +\infty$.

(i) *If* $\Lambda \in \mathcal{F}_\omega$, *there is a t-set* $S_0(\Lambda)$ (*independent of* M), *of measure* 0, *such that the sets*[23] $\{T_t\Lambda \cdot \mathrm{M}\}$, $t \in' S_0(\Lambda)$, *form a separable collection.*

(ii) *If* $\Lambda \in \mathcal{F}_\omega$, $\mu(T_t\Lambda \cdot \mathrm{M})$ *is a measurable function of t, for every* M.

(iii) *If* $\Lambda \in \mathcal{F}_\omega$ *and if* $\Phi_\Lambda^{\mathrm{M}}(t)$ *is the function, taking on values in the metric space* $\mathcal{F}^*$, *for which* $\Phi_\Lambda^{\mathrm{M}}(t_0)$ *corresponds to* $T_{t_0}\Lambda \cdot \mathrm{M}$, $\Phi_\Lambda^{\mathrm{M}}(t)$ *is a measurable function of t, for every* M.

(iv) *If* $\Lambda \in \mathcal{F}_\omega$, *there is a t-set* $S_0(\Lambda)$ *independent of* M, *of measure* 0, *such that, if* $t \in' S_0(\Lambda)$, $\mu(\mathrm{M} \cdot T_{t+h}\Lambda \cdot CT_t\Lambda) \to 0$ *when* $h \to 0$ *on a set (which may depend on t and* Λ *but not on* M) *having metric density* 1 *at* $h = 0$.

Conversely, if a family of set transformations under Case II *satisfies* (i) *and* (ii), *or* (i) *and* (iii), *or* (iv), *the family is measurable.*

If Λ_1 is a measurable ω-set, there is a set Λ_2 in $\mathcal{F}_\omega$ differing from Λ_1 by at most an ω-set of measure 0, so that, in Case II, $T_t\Lambda_2 \equiv T_t\Lambda_1$, and Λ can be taken as any ω-measurable set in the statements of Theorem 4. In (ii), the measurability of $\mu(T_t\Lambda \cdot \mathrm{M})$, for $\mu(\mathrm{M}) < \infty$, evidently implies the measurability of $\mu(T_t\Lambda \cdot \mathrm{M})$ for an arbitrary measurable ω-set M.

The family of set transformations is measurable if and only if to each ω-set Λ in $\mathcal{F}_\omega$ corresponds a measurable (t, ω)-set $\tilde{\Lambda}$ meeting $t = t_0$ in $T_{t_0}\Lambda$, if $T_{t_0}\Lambda$ is defined. Theorem 2 is then easily applied to obtain the stated results. Of the above examples, (d) is an example of a non-measurable family. In fact the collection of sets $\{T_t\Lambda_0\}$ discussed under (d), or any non-denumerable subcollection, was proved non-separable, and this contradicts (i). The other two examples in which an ω-measure was defined, (b) and (c), are examples of measurable families.

As in Theorem 2, there is an alternative statement to (iv) (in which M is to vary even if $\mu(\Omega) < +\infty$).

(iv') *If* $\Lambda \in \mathcal{F}_\omega$, $\mu(T_t\Lambda \cdot \mathrm{M})$ *is approximately continuous for almost all t. The exceptional t-set is independent of* M, *but may depend on* Λ.

THEOREM 5. *If* $\{T_t\}$ *is a measurable family of set transformations under Case* II, *then each* T_t *is defined on every measurable ω-set.*

Let Λ be a measurable ω-set. We shall show that $T_t\Lambda$ is defined for all values of t. According to Theorem 4, there is a t-set S_0, of measure 0, such that, if

[23] Throughout this paper, $T_t\Lambda \cdot \mathrm{M}$ will denote the set $(T_t\Lambda) \cdot \mathrm{M}$.

$t \epsilon' S_0$, $T_t\Lambda$ is defined, and the collection $\mathfrak{S}(\mathrm{M})$ of sets $\{T_t\Lambda \cdot \mathrm{M}\}$, $t \epsilon' S_0$, is a separable collection whenever M is a measurable ω-set of finite measure. It was seen above that there is then a t-set $S_1(\mathrm{M})$, of measure 0, such that T_t ($t \epsilon' S_1$) is defined on every set of $\mathfrak{S}(\mathrm{M})$. Now let $\mathrm{M}_1, \mathrm{M}_2, \cdots$ be a sequence of measurable ω-sets of finite measure whose sum is Ω, let $S_1 = \sum_1^\infty S_1(\mathrm{M}_n)$, and let t_0 be any value of t. Then there is a number τ such that $t_0 - \tau \epsilon' S_0$ and that $\tau \epsilon' S_1$. If τ is so chosen, $T_{t_0-\tau}\Lambda$ is defined, $T_{t_0-\tau}\Lambda \cdot \mathrm{M}_n$ is a set of $\mathfrak{S}(\mathrm{M}_n)$, and $T_\tau(T_{t_0-\tau}\Lambda \cdot \mathrm{M}_n)$ is defined. According to property (β) of the set transformations we are considering,

$$T_\tau\left(\sum_1^\infty T_{t_0-\tau}\Lambda \cdot \mathrm{M}_n\right) = T_\tau(T_{t_0-\tau}\Lambda)$$

is then defined, and $T_{t_0}\Lambda$ is also defined: $T_{t_0}\Lambda = T_\tau(T_{t_0-\tau}\Lambda)$, by property ($\alpha$).

THEOREM 6. *If the family of set transformations $\{T_t\}$ under Case* II *is measurable, there is a non-negative (t, ω)-measurable function $\varphi(t, \omega)$ (which may take on the value $+\infty$) such that if Λ is a measurable ω-set,*

$$\mu(T_t\Lambda) = \int_\Lambda \varphi(t, \omega)\, d\omega$$

for almost all t. If the measurable ω-sets form a separable collection, the exceptional t-set can be taken independent of Λ.

Let $\tilde{\Lambda}$ be any measurable (t, ω)-set, meeting $t = t_0$ in the ω-set Λ_{t_0}. Then Λ_t is ω-measurable for almost all t, so that (according to Theorem 5) $\mu(T_t\Lambda_t)$ is defined for almost all t. If $\tilde{\Lambda}$ is a product set $I \times \Lambda$, where I is a t-interval and Λ is a measurable ω-set, or if $\tilde{\Lambda}$ is a finite sum of such product sets, $\mu(T_t\Lambda_t)$ is evidently a measurable function of t. The function $\mu(T_t\Lambda_t)$ is then measurable for all (t, ω)-sets of the field $\mathfrak{F}_t \times \mathfrak{F}_\omega$, and therefore for all (t, ω)-measurable sets, since T_t takes ω-sets of measure 0 into ω-sets of measure 0. We define a new measure on the field of measurable (t, ω)-sets: if $\tilde{\Lambda}$ is in this field, the new measure of $\tilde{\Lambda}$ is defined as $\int_{-\infty}^{\infty} \mu(T_t\Lambda_t)\, dt$. This measure vanishes if $\tilde{\Lambda}$ is of (t, ω)-measure 0, since then $\mu(T_t\Lambda_t) = \mu(\Lambda_t) = 0$ for almost all t. Then by the Radon-Nikodym theorem,[24] there is a measurable (t, ω)-function $\varphi(t, \omega)$ such that

$$\int_{-\infty}^{\infty} \mu(T_t\Lambda_t)\, dt = \int\int_{\tilde{\Lambda}} \varphi(t, \omega)\, dt\, d\omega.$$

Let Λ be any measurable ω-set and let S_n be the t-set defined by the inequality $\mu(T_t\Lambda) \leq n$. Then if I is any t-interval, and if $\tilde{\Lambda}$ is the direct product $S_n \cdot I \times \Lambda$,

$$\int_{S_n \cdot I} \mu(T_t\Lambda)\, dt = \int_{S_n \cdot I} dt \int_\Lambda \varphi(t, \omega)\, d\omega$$

[24] S. Saks, *Theory of the Integral*, New York, 1937, p. 36. Saks proves the theorem for finite-valued set functions, but the extension to the present case can be made without difficulty.

so that

$$\mu(T_t\Lambda) = \int_\Lambda \varphi(t, \omega)\, d\omega$$

almost everywhere on S_n. Letting n become infinite, we find that the representation of $\mu(T_t\Lambda)$ is valid for almost all t in $\sum_1^\infty S_n$. If $\mu(T_t\Lambda) = +\infty$ on a set S_∞ of positive measure, we find similarly that

$$\mu(T_t\Lambda) = \int_\Lambda \varphi(t, \omega)\, d\omega = +\infty$$

almost everywhere on S_∞. If a denumerable collection of ω-sets is preassigned, the exceptional t-set can be taken independent of Λ in this collection, since a denumerable sum of sets of measure 0 has measure 0. It follows at once that if the measurable ω-sets form a separable collection, the exceptional t-set can be taken independent of Λ.

Theorem 6, under the hypothesis that the measurable ω-sets form a separable collection, can also be derived if we use Theorem 1 and the fact (which follows at once by applying the Radon-Nikodym theorem to the set function $\mu(T_t\Lambda)$) that for each value of t there is an ω-measurable function $\varphi_0(t, \omega)$ such that

$$\mu(T_t\Lambda) = \int_\Lambda \varphi_0(t, \omega)\, d\omega$$

for all measurable ω-sets Λ.

LEMMA 3. *If E is any measurable t-set of positive finite measure, and if $\{\delta_n\}$ is a sequence of numbers converging to* 0, *there is a subset E' of E, of measure arbitrarily near that of E, and a subsequence $\{\delta_{a_n}\}$ of $\{\delta_n\}$ such that, if $t \in E'$, $t + \delta_{a_n} \in E$ $(n = 1, 2, \cdots)$.*

By a theorem of Auerbach,[25] if $\psi(t)$ is the characteristic function of E, there is a subsequence $\{\delta_{a_n}\}$ of $\{\delta_n\}$ such that $\psi(t + \delta_{a_n}) \to \psi(t)$ almost everywhere. According to Egoroff's theorem, there is a subset E' of E of measure arbitrarily near that of E and such that $\psi(t + \delta_{a_n})$ converges uniformly to $\psi(t)$ for $t \in E'$. Then there is an integer N such that $\psi(t + \delta_{a_n}) > \frac{1}{2}$ (so that $\psi(t + \delta_{a_n}) = 1$) for $t \in E'$, if $n > N$. The sequence $\{\delta_{a_n}\}$ can be taken as $\{\delta_{a_{N+n}}\}$.

THEOREM 7. *Let $\{T_t\}$ be a family of set transformations under Case* II. *Then $\mu(T_t\Lambda \cdot \mathrm{M})$ is a continuous function of t (whenever Λ, M are measurable ω-sets and $\mu(\mathrm{M})$ is finite) if and only if* (a) *the family is measurable and* (b) *$\mu(T_t\Lambda \cdot \mathrm{M})$ is uniformly small (in t for each fixed M) with $\mu(\Lambda \cdot \mathrm{M})$ for t in every finite t-interval.*

It is sufficient above if the family is measurable and if there is a measurable t-set E of positive measure such that whenever $\mu(\mathrm{M}) < \infty$, $\mu(T_t\Lambda \cdot \mathrm{M})$ is uniformly small (in t for each fixed M) with $\mu(\Lambda \cdot \mathrm{M})$ for t in E.

The continuity of $\mu(T_t\Lambda \cdot \mathrm{M})$, for measure preserving transformations, follows readily from a theorem of von Neumann on one-parameter families of unitary transformations in Hilbert space.[26] In proving that the given condition is

[25] Fundamenta Mathematicae, vol. 11(1923), pp. 196–197.

[26] Annals of Mathematics, vol. 33(1932), pp. 568–569.

sufficient, if $\mu(\Omega) < \infty$, we can replace the condition that $\mu(T_t\Lambda\cdot\mathrm{M})$ be small with $\mu(\Lambda\cdot\mathrm{M})$ for each M by the condition that $\mu(T_t\Lambda)$ be small with $\mu(\Lambda)$ (uniformly in the t-set E).

If the family is measurable, and if there is a t-set E as described, let Λ, M be measurable ω-sets with $\mu(\mathrm{M}) < \infty$, and let $\Phi_\Lambda^{\mathrm{M}}(t)$ be the $\mathfrak{F}^*$-point corresponding to $T_t\Lambda\cdot\mathrm{M}$ for each value of t. Then it will be sufficient to show that $\Phi_\Lambda^{\mathrm{M}}(t)$ is a continuous function of t. We shall do this for a fixed pair: $\Lambda = \Lambda'$, $\mathrm{M} = \mathrm{M}'$.[27] It follows from Lemma 1 and Theorem 4 that every $\Phi_\Lambda^{\mathrm{M}}(t)$ is approximately continuous for almost all t. Let $\mathrm{M}_1 = \mathrm{M}'$, M_2, $\cdots$ be a sequence of measurable ω-sets of finite measure whose sum is Ω. According to Theorem 4(i) there is a t-set S_n of measure 0 such that the sets $\{T_t\Lambda'\cdot\mathrm{M}_n\}$, $t \epsilon' S_n$, form a separable collection $\mathfrak{S}_n$. Let

$$S_0 = \sum_1^\infty S_n, \qquad \mathfrak{S} = \sum_1^\infty \mathfrak{S}_n,$$

and let $\Lambda_1, \Lambda_2, \cdots$ be a sequence of measurable ω-sets everywhere dense (in terms of $\mathfrak{F}^*$-topology) on $\mathfrak{S}$. Let I be a t-interval such that $m(E\cdot I) > 0$. According to Lemma 1(ii) (Lusin's theorem), there is a measurable subset E_{jk} of $E\cdot I$, of measure at least

$$[1 - 2^{-j-k-1}]m(E\cdot I),$$

such that $\Phi_{\Lambda_j}^{\mathrm{M}_k}(t)$ is continuous on E_{jk}, relative to E_{jk}. If F is defined as $CS_0\cdot\prod_{j,k=1}^\infty E_{jk}$, $mF \geqq \frac{1}{2}m(E\cdot I)$ and $\Phi_{\Lambda_j}^{\mathrm{M}_k}(t)$ is continuous on F, relative to F, for all j, k. If Λ is a set in $\mathfrak{S}$, there is a subsequence of sets $\{\Lambda_{a_n}\}$ such that $\Lambda_{a_n} \to \Lambda$ in the sense of $\mathfrak{F}^*$-topology. Then $\Phi_{\Lambda_{a_n}}^{\mathrm{M}_k}(t) \to \Phi_\Lambda^{\mathrm{M}_k}(t)$ uniformly on E, according to the hypotheses of Theorem 7, so that $\Phi_\Lambda^{\mathrm{M}_k}(t)$ is continuous on F, relative to F. The same reasoning shows that $\Phi_\Lambda^{\mathrm{M}_k}(t)$ is continuous on F, relative to F, whenever Λ is a finite or denumerably infinite sum of sets in $\mathfrak{S}$—for example, if $\Lambda = T_t\Lambda'$, $t \epsilon' S_0$. We shall now show that if t_0 is any real number, $\Phi_{\Lambda'}^{\mathrm{M}'}(t)$ is continuous at t_0. To show this, it will be sufficient to show that, whenever $\{\delta_n\}$ is a sequence of numbers converging to 0, there is a subsequence $\{\delta_{\alpha_n}\}$ for which $\Phi_{\Lambda'}^{\mathrm{M}'}(t_0 + \delta_{\alpha_n}) \to \Phi_{\Lambda'}^{\mathrm{M}'}(t_0)$. By Lemma 3, if $\{\delta_n\}$ is a sequence of numbers converging to 0, there is a subsequence $\{\delta_{\alpha_n}\}$ and a subset F' of F, of positive measure, such that, if $t \epsilon F'$, $t + \delta_{\alpha_n} \epsilon F$ $(n = 1, 2, \cdots)$. Now let τ be a number in F' such that $t_0 - \tau \epsilon' S_0$, and let Λ'' be the set $T_{t_0-\tau}\Lambda'$. Then

$$\Phi_{\Lambda'}^{\mathrm{M}'}(t_0) = \Phi_{\Lambda''}^{\mathrm{M}'}(\tau), \qquad \Phi_{\Lambda'}^{\mathrm{M}'}(t_0 + \delta_{\alpha_n}) = \Phi_{\Lambda''}^{\mathrm{M}'}(\tau + \delta_{\alpha_n}).$$

But $\tau, \tau + \delta_{\alpha_1}, \tau + \delta_{\alpha_2}, \cdots$ are all in F, so (remembering that $\mathrm{M}' = \mathrm{M}_1$)

$$\Phi_{\Lambda'}^{\mathrm{M}'}(t_0 + \delta_{\alpha_n}) = \Phi_{\Lambda''}^{\mathrm{M}'}(\tau + \delta_{\alpha_n}) \to \Phi_{\Lambda''}^{\mathrm{M}'}(\tau) = \Phi_{\Lambda'}^{\mathrm{M}'}(t_0),$$

as was to be proved.

[27] If $\mu(\Omega) < \infty$, only the case $\mathrm{M}' = \Omega$ need be considered, and the proof can be considerably simplified.

Conversely, if $\mu(T_t\Lambda\cdot M)$ is continuous whenever $\mu(M) < \infty$, the family of set transformations is measurable, by Theorem 4(iv′). Let M be a fixed measurable ω-set of finite measure. Then for each value of t, $\mu(T_t\Lambda\cdot M)$ is a finite-valued absolutely continuous function of sets Λ, and therefore approaches 0 with $\mu(\Lambda)$. Unless $\mu(T_t\Lambda\cdot M)$ is uniformly small with $\mu(\Lambda\cdot M)$ for t in every finite interval, there is a finite interval I, a sequence $\{t_n\}$ of values of t in I, a sequence of measurable ω-sets $\Lambda_1, \Lambda_2, \cdots$, and a positive number ϵ such that

$$\mu(\Lambda_n) \to 0, \qquad \mu(T_{t_n}\Lambda_n\cdot M) \geqq \epsilon, \qquad (n = 1, 2, \cdots).$$

We can suppose subsequences have been chosen, if necessary, so that the sequence $\{t_n\}$ converges: $t_n \to t_0$, and that the numbers $\{\mu(T_{t_0}\Lambda_n\cdot M)\}$ which we have already shown converge to 0 with n^{-1} form a convergent sum:

$$\sum_1^\infty \mu(T_{t_0}\Lambda_n\cdot M) < \infty.$$

Then if $\Lambda^N = \sum_N^\infty \Lambda_n$,

$$\epsilon \leqq \mu(T_{t_n}\Lambda_n\cdot M) \leqq \mu(T_{t_n}\Lambda^N\cdot M) \qquad (n \geqq N),$$

so that, letting n become infinite, and using the continuity of $\mu(T_t\Lambda\cdot M)$, we get

$$\epsilon \leqq \mu(T_{t_0}\Lambda^N\cdot M) \leqq \sum_N^\infty \mu(T_{t_0}\Lambda_n\cdot M).$$

This inequality is impossible for large values of N, since the sum on the right converges to 0 with N^{-1}. Therefore $\mu(T_t\Lambda\cdot M)$ is uniformly small with $\mu(\Lambda\cdot M)$, $t \in I$, as was to be proved.

In Example (a) above, we did not define an ω-measure. If we choose some finite set of points on the perimeter of the circle and define the measure of an ω-set as the number of these points in it, $\mu(T_t\Lambda\cdot M)$ is not a continuous function of t, if Λ, M are suitably chosen. In Example (b), Theorem 7 may or may not be applicable, depending on the choice of $f(x, y)$. In Example (c) Theorem 7 is applicable. We have already seen that the family $\{T_t\}$ of Example (d) is not measurable.

THEOREM 8. *Let $\{T_t\}$ be a measurable family of set transformations (under Case I or II), and suppose that there is a denumerable collection $\mathfrak{S}$ of measurable ω-sets such that if $\mathfrak{S}_*$ is the collection of sets which are finite or denumerably infinite sums of sets of $\mathfrak{S}$, each set of $\mathfrak{S}_*$ is taken by T_t into a set of $\mathfrak{S}_*$ for almost all values of t. Then $\mu(T_t\Lambda\cdot M)$ is a lower semi-continuous function of t if $\Lambda \in \mathfrak{S}_*$, and if M is a measurable ω-set of finite measure.*

Let M be any ω-measurable set of finite measure. Applying Lusin's theorem, we see that there is a measurable t-set F of positive measure such that $\mu(T_t\Lambda\cdot M)$ is continuous on F relative to F, for Λ in $\mathfrak{S}$ or a finite sum of sets in $\mathfrak{S}$. Now let Λ be a set in $\mathfrak{S}_*$:

$$\Lambda = \sum_1^\infty \Lambda_n, \qquad \Lambda_n \in \mathfrak{S}, \qquad (n = 1, 2, \cdots).$$

Then

$$\mu(T_t\Lambda\cdot\mathrm{M}) = \lim_{N\to\infty}\mu\left(\sum_1^N T_t\Lambda_n\cdot\mathrm{M}\right).$$

The limit is that of a monotone non-decreasing sequence of functions continuous on F, relative to F. Then $\mu(T_t\Lambda\cdot\mathrm{M})$ is lower semi-continuous on F, relative to F. A modification of the reasoning used in the proof of Theorem 7 now completes the proof of the present theorem.

Probably the most important application of Theorem 8 is the following. Let Ω be a separable topological space, and suppose that open sets are measurable. If the set transformations take open sets into open sets, the hypotheses of Theorem 8 are satisfied, and $\mu(T_t\Lambda\cdot\mathrm{M})$ will be lower semi-continuous for Λ open and M measurable and of finite measure.

THEOREM 9. *Let* $\{T_t\}$ *be a measurable family of set transformations under Case* I, *generated by point transformations. Suppose that* Ω *is a separable topological Hausdorff space and that* $\mathcal{F}_\omega$ *is the Borel field determined by the open sets of* Ω. *Suppose that there is a t-set* S_0 *of measure* 0 *such that, whenever* $t_n \to t \,\epsilon'\, S_0$, $T_{t_n}\omega \to T_t\omega$ *almost everywhere on* Ω. *Then, if we neglect an* ω*-set of measure* 0, *each trajectory* $\{T_t\omega\}$ $(-\infty < t < \infty)$ *is continuous relative to a t-set* (*depending on* ω) *whose complement is of measure* 0.

The function $T_t\omega$ of (t, ω) is a function taking on values in the separable Hausdorff space Ω. Let Λ be an open set of Ω, and let $\varphi(\omega)$ be the characteristic function of Λ. Then $\varphi(\omega)$ is measurable with respect to the field $\mathcal{F}_\omega$ so $\varphi(T_t\omega)$ is (t, ω)-measurable, i.e., the (t, ω)-set for which $T_t\omega \,\epsilon\, \Lambda$ is measurable. The function $T_t\omega$ is thus (t, ω)-measurable. The function $T_t\omega = f(t, \omega)$ satisfies the hypotheses of Theorem 3, and the conclusion of that theorem becomes the conclusion of Theorem 9.

Let Ω, Ω' be two abstract spaces, with ω, ω' measures as described above. Let $\{T_t\}$, $\{T_t'\}$ be families of set transformations, under Case II, applied to the measurable sets of Ω, Ω', respectively. Then these two families will be called *equivalent* if there is a transformation $\Lambda' = \mathfrak{T}\Lambda$ of the measurable ω-sets Λ into the measurable ω'-sets Λ' with the following four properties.

(i) The transformation $\mathfrak{T}$, defined on every measurable ω-set, is multiple-valued: two measurable ω'-sets are images of the same ω-set if and only if they differ by at most an ω'-set of measure 0. The range of $\mathfrak{T}$ includes every measurable ω'-set.

(ii) $\mu'(\mathfrak{T}\Lambda) = \mu(\Lambda)$.

These two conditions imply that $\mathfrak{T}$ induces a one-to-one continuous correspondence between the topological spaces $\mathcal{F}^*$, $\mathcal{F}'^*$.

(iii) If Λ_1, Λ_2, $\cdots$ is any sequence of measurable ω-sets,

$$\mathfrak{T}\left(\sum_1^\infty \Lambda_j\right) = \sum_1^\infty \mathfrak{T}\Lambda_j, \qquad \mathfrak{T}\prod_1^\infty \Lambda_j = \prod_1^\infty \mathfrak{T}\Lambda_j.$$

(iv) If $T_t\Lambda$ and $T'_t\mathfrak{T}\Lambda$ are defined, then $\mathfrak{T}T_t\Lambda = T'_t\mathfrak{T}\Lambda$. Evidently if two families of set transformations are equivalent, either neither is measurable or both are. The equivalence relation is reflexive, symmetric, and transitive.

THEOREM 10. *Let $\{T_t\}$ be a family of set transformations under Case II. There is then an equivalent family of set transformations $\{T'_t\}$ operating on sets of a space Ω', with the following properties: there is a uniquely defined point transformation $\overline{T}'_t\omega'$ defined on Ω', for which $\overline{T}'_{s+t}\omega' = \overline{T}'_s(\overline{T}'_t\omega')$ identically in s, t, ω', and such that if Λ' is a measurable ω'-set, $\overline{T}'_t\Lambda'$ is one of the images of Λ' under T'_t.*

Less precisely, the theorem states that the given family is equivalent to a family generated by point transformations. With different hypotheses, von Neumann has proved[28] that a given set transformation of a certain type can always be considered as generated by a point transformation. It has been seen above that this is not true of the set transformations being considered here.

Let Ω' be the space of single-valued functions $\xi(t)$, defined for $-\infty < t < \infty$ and taking on values in Ω. A point ω' will then be a function $\xi(t)$. Let n be any positive integer, let $t_1, \cdots, t_n$ be any n distinct values of t, and let $\Lambda_1, \cdots, \Lambda_n$ be any n sets in $\mathfrak{F}_\omega$. We shall call the ω'-set[29] $\{\xi(t_j) \in \Lambda_j \; (j = 1, \cdots, n)\}$ an elementary ω'-set. Let $\mathfrak{T}'\Lambda'$ be the (multiple-valued) set transformation, defined for Λ' any elementary ω'-set and taking $\{\xi(t_j) \in \Lambda_j \; (j = 1, \cdots, n)\}$ into every ω-set $\prod_1^n T_{-t_j}\Lambda_j$. If

$$\{\xi(t_j) \in \Lambda_j \quad (j = 1, \cdots, n)\} = \{\xi(t_j^0) \in \Lambda_j^0 \quad (j = 1, \cdots, n_0)\},$$

the condition on $\xi(t)$ for a t-value τ not present in both $t_1, \cdots, t_n, t_1^0, \cdots, t_{n_0}^0$ is simply $\xi(\tau) \in \Omega$, while if $\tau = t_j = t_k^0$, $\Lambda_{t_j} = \Lambda_{t_k^0}^0$. Then $\mathfrak{T}'\Lambda'$ depends only on Λ', not on the particular t-values used in defining Λ'. Two different images of Λ' differ by at most an ω-set of measure 0. We shall extend the domain of definition of $\mathfrak{T}'$, keeping the property that two images under $\mathfrak{T}'$ of a set Λ' can differ by at most an ω-set of measure 0.

Suppose that an elementary set can be expressed as a finite sum of elementary sets:[30]

$$\text{(A)} \qquad \{\xi(t_j) \in \Lambda_j \quad (j = 1, \cdots, n)\} = \sum_{k=1}^{\nu} \{\xi(t_j) \in \Lambda_j^k \quad (j = 1, \cdots, n)\}.$$

Then we shall show that the transformation $\mathfrak{T}'$ is additive:

$$\text{(B)} \qquad \prod_1^n T_{-t_j}\Lambda_j = \sum_{k=1}^{\nu} \prod_{j=1}^{n} T_{-t_j}\Lambda_j^k.$$

[28] Annals of Mathematics, vol. 33(1932), p. 582.

[29] The notation $\{\cdots\}$ will be used to denote the set determined by the conditions written between the braces.

[30] Since only a finite number of sets is involved, we can suppose that the same t-values, $t_1, \cdots, t_n$, are involved in the definitions of all the sets, supposing that some Λ_j or Λ_j^k sets are Ω, if necessary.

(Throughout this section we shall neglect ω-sets of measure 0.) If $n = 1$, (A) implies (B) because of the additivity of T_t. The proof for $n = 2$ is characteristic of the proof for $n > 1$, and only the case $n = 2$ will be considered. By operations on the Λ_1^k not affecting the validity of (A), and which do not lead to changes in $\sum_{k=1}^{\nu} \prod_{j=1}^{n} T_{-t_j}\Lambda_j^k$, we can bring it about that two Λ_1^k-sets are either disjunct or identical. In the latter case we can combine two terms of the right side of (A), again without changing the right side of (B), so that finally the Λ_1^k-sets are all disjunct (and we can suppose that none are empty). But then (A) implies that

$$\sum_{k=1}^{\nu} \Lambda_1^k = \Lambda_1, \qquad \Lambda_2^k = \Lambda_2, \qquad (k = 1, \cdots, \nu).$$

Equation (B) follows from these facts; for

$$\sum_{k=1}^{\nu} \prod_{j=1}^{2} T_{-t_j}\Lambda_j^k = T_{-t_2}\Lambda_2 \sum_{k=1}^{\nu} T_{-t_1}\Lambda_1^k = \prod_{1}^{2} T_{-t_j}\Lambda_j .$$

It follows at once from this result that if

$$\text{(C)} \qquad \sum_{k=1}^{l} \{\xi(t_j) \in \Lambda_j^k \quad (j = 1, \cdots, n)\} = \sum_{k=1}^{m} \{\xi(t_j) \in \mathrm{M}_j^k \quad (j = 1, \cdots, n)\},$$

then

$$\text{(D)} \qquad \sum_{k=1}^{l} \prod_{j=1}^{n} T_{-t_j}\Lambda_j^k = \sum_{k=1}^{m} \prod_{j=1}^{n} T_{-t_j}\mathrm{M}_j^k,$$

if we neglect sets of measure 0. Then if an ω'-set Λ' is a finite sum of elementary sets (say if Λ' is the left side of (C)), we can define $\mathfrak{T}\Lambda'$ (as the left side of (D)) and thus obtain an unambiguous definition of $\mathfrak{T}\Lambda'$, aside from ω-sets of measure 0. The complement of an elementary set is a finite sum of elementary sets. The collection of finite sums of elementary sets then forms a field, and the transformation $\mathfrak{T}'$ has been defined on this field and is additive on it.

Now let Λ' be an elementary ω'-set which is a sum of denumerably many elementary ω'-sets: $\Lambda' = \sum_{1}^{\infty} \Lambda_n$. Then $\Lambda' - \sum_{1}^{\nu} \Lambda_n' = \mathrm{M}_\nu'$ is itself a sum of a finite number of elementary sets, so that

$$\mathfrak{T}'\Lambda' = \sum_{1}^{n} \mathfrak{T}'\Lambda_j' + \mathfrak{T}'\mathrm{M}_n'.$$

We shall show $\mathfrak{T}'\Lambda' = \sum_{1}^{\infty} \mathfrak{T}'\Lambda_n'$ by showing that $\lim_{n\to\infty} \mu(\mathfrak{T}'\mathrm{M}_n') = 0$. The sequence $\mathrm{M}_1', \mathrm{M}_2', \cdots$ is monotone non-increasing, and $\prod_{1}^{\infty} \mathrm{M}_n' = 0$. Let $t_1, t_2, \cdots$ be

the t-values involved in defining the elementary sets of M_1', M_2', $\cdots$ and let M_ν' be the set

$$\sum_{k=1}^{m_\nu} \{\xi(t_j) \in M_j^k(\nu)\ (j \geqq 1)\},$$

where $M_j^k(\nu) = \Omega$ for large j, $j > j(\nu, k)$. By definition,

$$\mathfrak{T}'M_\nu' = \sum_{k=1}^{m_\nu} \prod_{j=1}^{\infty} T_{-t_j} M_j^k(\nu).$$

The sets $\mathfrak{T}'M_1'$, $\mathfrak{T}'M_2'$, $\cdots$ form a monotone non-increasing sequence (if we neglect ω'-sets of measure 0); let M be their intersection. Then, neglecting ω-sets of measure 0, we have

$$\sum_{k=1}^{m_\nu} \prod_{j=1}^{\infty} T_{-t_j} M_j^k(\nu) \supseteq M \qquad (\nu = 1, 2, \cdots),$$

and we must show that $\mu(M) = 0$. Define $M^k(\nu)$ by:

$$M^k(\nu) = \prod_{j=1}^{\infty} T_{-t_j} M_j^k(\nu),$$

so that we have

$$\sum_{k=1}^{m_\nu} M^k(\nu) \supseteq M \qquad (\nu \geqq 1),$$

$$M^k(\nu) \subseteq T_{-t_j} M_j^k(\nu), \qquad T_{t_j} M^k(\nu) \subseteq M_j^k(\nu), \qquad (j, k, \nu \geqq 1),$$

neglecting ω-sets of measure 0. Then

$$\sum_{k=1}^{m_\nu} \{\xi(t_j) \in T_{t_j} M^k(\nu)\ (j \geqq 1)\} \subseteq M_\nu' \to 0,$$

$$\sum_{k=1}^{m_\nu} M^k(\nu) \supseteq M \qquad (\nu \geqq 1),$$

and we shall derive a contradiction from these two relations, under the hypothesis that $\mu(M) > 0$. Suppose first that Ω is the x-interval $0 \leqq x \leqq 1$ and that the $M_j^k(\nu)$-sets are Borel sets. Then μ-measure on these sets, and on the field they determine, is a measure of Borel sets, so there is a closed bounded subset $\bar{M}_j^k(\nu)$ of $T_{t_j}M^k(\nu)$ with $\bar{M}_j^k(\nu) = \Omega$, if $M^k(\nu) = \Omega$, such that $\mu(T_{-t_j}\bar{M}_j^k(\nu))$ is so close to $\mu(M^k(\nu))$ that, if we set $\bar{M}^k(\nu) = \prod_{j=1}^{\infty} T_{-t_j}\bar{M}_j^k(\nu)$,

$$\mu(M - M \sum_{k=1}^{m_\nu} \bar{M}^k(\nu)) \leqq \mu\left[\sum_{k=1}^{m_\nu} (M^k(\nu) - \bar{M}^k(\nu))\right] < \mu(M)3^{-\nu} \quad (\nu \geqq 1).$$

We have

$$T_{t_j} \bar{M}^k(\nu) \subseteq \bar{M}_j^k(\nu) \subseteq T_{t_j} M^k(\nu) \qquad (j, k, \nu \geqq 1).$$

Define $\bar{M}'_n \subseteq M'_n$ by

$$\bar{M}'_n = \prod_{\nu=1}^{n} \sum_{k=1}^{m_\nu} \{\xi(t_j) \epsilon \bar{M}_j^k(\nu) \ (j \geqq 1)\}.$$

Then

$$\bar{M}'_q \supseteq \prod_{\nu=1}^{q} \sum_{k=1}^{m_\nu} \{\xi(t_j) \epsilon T_{t_j} \bar{M}^k(\nu) \ (j \geqq 1)\} = \sum_{l=1}^{n_q} \{\xi(t_j) \epsilon T_{t_j} \bar{M}^l(q) \ (j \geqq 1)\},$$

where $\bar{M}^l(q)$ (q fixed) runs through all products $\bar{M}^{i_1}(1) \cdot \bar{M}^{i_2}(2) \cdots \bar{M}^{i_q}(q)$ $(1 \leqq i_j \leqq m_j)$. Denote the set on the right by $\bar{\bar{M}}'_q$. For each q,

$$\mu\left(\sum_{l=1}^{n_q} \bar{M}^l(q)\right) = \mu\left(\prod_{j=1}^{q} \sum_{k=1}^{m_j} \bar{M}^k(j)\right) > \mu(M)(1 - 3^{-1} - 3^{-2} - \cdots - 3^{-q})$$
$$> \tfrac{1}{2}\mu(M),$$

so that some $\bar{M}^l(q)$ and therefore some $T_{t_j}\bar{M}^l(q)$ is of positive measure. Then $\bar{\bar{M}}'_q$ is not empty for any q, and $\bar{\bar{M}}'_1 \supseteq \bar{\bar{M}}'_2 \supseteq \cdots \to 0$. Let $\omega_n : \xi_n(t)$ be a point of $\bar{\bar{M}}'_n$. Then by the usual diagonal procedure, there is an increasing sequence of integers $a_1, a_2, \cdots$ such that $\lim_{n \to \infty} \xi_{a_n}(t_j) = \xi_j$ exists for all j. Any point $\omega : \xi(t)$ with $\xi(t_j) = \xi_j$ $(j \geqq 1)$ is in $\bar{M}'_n$ for every n, since all the $\bar{M}_j^k$ are closed, and this contradicts the fact that $\prod_1^\infty \bar{M}'_n = 0$. This proves the result desired in the special case being considered: Ω the interval $0 \leqq x \leqq 1$, $M_j^k(\nu)$ Borel. We reduce the general case to this as follows. There is a collection of ω-sets $\{M_r\}$, defined for r rational, $0 < r < 1$, and satisfying the following conditions.

1. Each M_r is in the field of sets determined by the $M_j^k(\nu)$, and conversely.
2. If $r_1 < r_2$, $M_{r_1} \subseteq M_{r_2}$.
3. If $r_1 \to r_2$ from above, r_n, r rational, $\prod_{n=1}^{\infty} M_{r_n} = M_r$;[31]

$$\prod_{r>0} M_r = 0, \qquad \sum_{r<1} M_r = \sum_{j,k,\nu} M_j^k(\nu).$$

To each point ω we make correspond the real number x, $0 \leqq x \leqq 1$, satisfying

$$x = \underset{\omega \epsilon' M_r}{\text{G.L.B.}}\ r.$$

Then every x-value corresponds to a certain ω-set, perhaps the null set. The interval $0 < x \leqq r$ corresponds to M_r, if r is rational, so that a measure is determined on the Borel sets of $0 \leqq x \leqq 1$ by taking as the measure of a Borel set the ω-measure of the corresponding ω-set, which is necessarily in the Borel field of sets determined by the $M_j^k(\nu)$. The general case is thus reduced to the case just proved.

We have now shown that (if all the sets involved are elementary sets) Λ'

[31] Transactions of the American Mathematical Society, vol. 44(1938), p. 91.

$= \sum_1^\infty \Lambda_n'$ implies that $\mathfrak{T}'\Lambda' = \sum_1^\infty \mathfrak{T}'\Lambda_n'$. It follows at once from this that if an ω'-set can be expressed in two ways as a finite or denumerably infinite sum of elementary sets:

$$\sum_1^\infty \Lambda_n' = \sum_1^\infty \mathrm{M}_n',$$

then

$$\sum_1^\infty \mathfrak{T}'\Lambda_n' = \sum_1^\infty \mathfrak{T}'\mathrm{M}_n',$$

so that if Λ' is a set which can be expressed as a finite or denumerably infinite sum of elementary sets, $\Lambda' = \sum_1^\infty \Lambda_n'$, we can define $\mathfrak{T}'\Lambda'$ unambiguously (neglecting sets of measure 0) as $\sum_1^\infty \mathfrak{T}'\Lambda_n'$. We now define an ω'-measure. In the following we shall assume throughout that $\mu(\Omega) < \infty$. The case $\mu(\Omega) = \infty$ requires unessential modifications. Let Λ' be any set in the field of finite sums of elementary ω'-sets. Then we define $\mu'(\Lambda')$ by

$$\mu'(\Lambda') = \mu(\mathfrak{T}'\Lambda').$$

From the properties of $\mathfrak{T}'$ and of ω-measure, this ω'-measure is completely additive on its field of definition, so the field of definition can be extended to the Borel field of sets determined by the elementary sets, and the measure of any measurable ω'-set Λ' is the limit of the measures of a monotone non-increasing sequence of sets, containing Λ', which are denumerable sums of elementary sets:[32]

$$\Lambda_1' \supseteq \Lambda_2' \supseteq \cdots \to \Lambda', \qquad \mu'(\Lambda') = \lim_{n\to\infty} \mu'(\Lambda_n').$$

We now define $\mathfrak{T}'\Lambda'$ as $\prod_1^\infty \mathfrak{T}'\Lambda_n'$. This definition of $\mathfrak{T}'\Lambda'$ is easily verified to be uniquely determined by Λ' and to be consistent with previous definitions if $\mathfrak{T}'\Lambda'$ is already defined. The transformation $\mathfrak{T}'$ is now defined on every ω'-measurable set and takes it into ω-measurable sets having the same measure. There is an inverse $\mathfrak{T}$, defined on every ω-measurable set and taking it into ω'-measurable sets of the same measure. The transformation $\mathfrak{T}$ takes the measurable ω-set Λ into every ω'-set

$$\mathfrak{T}\Lambda = \{\xi(t) \epsilon T_t\Lambda\}$$

(t fixed), any two such sets differing by at most an ω'-set of measure 0. Then, neglecting ω-sets of measure 0 (putting $t = -h$ and replacing Λ by $T_h\Lambda$ in the above equation), we get

$$\mathfrak{T}T_h\Lambda = \{\xi(-h) \epsilon \Lambda\}.$$

[32] H. Hahn, Pisa Annali, vol. 2(1933), p. 433.

Now let $\overline{T}'_h$ be the point transformation taking $\omega':\xi(t)$ into $\overline{T}'_h\omega':\xi(t + h)$. Then

$$\overline{T}'_h\mathfrak{T}\Lambda = \overline{T}'_h\{\xi(0) \epsilon \Lambda\} = \{\xi(-h) \epsilon \Lambda\},$$

if we neglect ω'-sets of measure 0, i.e.,

$$\mathfrak{T}T_h\Lambda = \overline{T}'_h\mathfrak{T}\Lambda.$$

The set transformation T'_t of Theorem 10 is that generated by the point transformation $\overline{T}'_t$, if we neglect ω'-sets of measure 0.

UNIVERSITY OF ILLINOIS.

SUFFICIENT CONDITIONS FOR THE CONVERGENCE OF A CONTINUED FRACTION

By Walter Leighton

1. **A new convergence criterion.** Continued fractions of the form

$$(1.1) \qquad 1 + \frac{a_1}{1+} \frac{a_2}{1+} \cdots \qquad (a_n \neq 0),$$

where the quantities a_n are arbitrary complex numbers, are of particular importance from a function-theoretic point of view.[1] J. Worpitsky, E. B. Van Vleck, and A. Pringsheim proved independently that the conditions $|a_n| \leqq \frac{1}{4}$ $(n = 2, 3, \cdots)$ are sufficient to insure the convergence of (1.1).[2] O. Szász [2] showed that $\frac{1}{4}$ was the best such constant by pointing out that the continued fraction (1.1) with

$$a_n = -\tfrac{1}{4} - \epsilon \qquad (n = 1, 2, 3, \cdots)$$

diverges for every $\epsilon > 0$. Szász [1] proved that a sufficient condition for the convergence of (1.1) is

$$\sum_{n=2}^{\infty} |a_n| - \sum_{n=2}^{\infty} R(a_n) < 2,$$

where $R(a_n)$ is the real part of a_n. Leighton and Wall [1] proved that the conditions $|a_{2n+1}| \leqq \frac{1}{4}$, $|a_{2n}| \geqq \frac{25}{4}$ $(n = 1, 2, 3, \cdots)$ are sufficient. All the above conditions require at least an infinite subsequence of the numbers a_n to be $\leqq \frac{1}{4}$ in absolute value. The following theorem removes this condition in a rather unexpected manner.

Theorem. *If the numbers a_n satisfy the conditions*

$$|1 + a_2| \geqq |a_1| + 1, \qquad |a_3| \geqq \frac{2+m}{1-m},$$

$$|a_{2n}| \leqq m \qquad (n = 2, 3, 4, \cdots),$$

$$|a_{2n+1}| \geqq 2 + m + m\,|a_{2n-1}| \qquad (n = 2, 3, 4, \cdots),$$

where m is any positive number < 1, the continued fraction (1.1) *converges.*

Received June 3, 1938.

[1] See, for example, Perron [2], Chapter VIII. (Numbers in brackets refer to the bibliography at the end of the paper.) W. T. Scott, in preparing a Rice Institute thesis, has found recently a number of results which strengthen significantly a natural generalization (Leighton [1]) of the material discussed by Perron (loc. cit.).

[2] O. Szász [2] discusses the history of this criterion.

To prove the theorem let us recall that the n-th approximant A_n/B_n of a continued fraction

$$b_0 + \frac{a_1}{b_1+}\,\frac{a_2}{b_2+}\cdots, \tag{1.1)'$$

where the a_n and b_n are arbitrary complex numbers, is defined by the following recursion relations:

$$\begin{aligned} A_0 &= b_0, & B_0 &= 1, \\ A_1 &= b_0b_0 + a_1, & B_1 &= b_1, \\ A_n &= b_nA_{n-1} + a_nA_{n-2}, & B_n &= b_nB_{n-1} + a_nB_{n-2} \end{aligned} \tag{1.2}$$

$$(n = 2, 3, 4, \cdots).$$

Associated with (1.1) are the two continued fractions

$$(1 + a_1) + \frac{-a_1a_2}{(1 + a_2 + a_3)+}\,\frac{-a_3a_4}{(1 + a_4 + a_5)+}\,\frac{-a_5a_6}{(1 + a_6 + a_7)+}\cdots, \tag{1.3}$$

$$1 + \frac{a_1}{1 + a_2+}\,\frac{-a_2a_3}{(1 + a_3 + a_4)+}\,\frac{-a_4a_5}{(1 + a_5 + a_6)+}\cdots. \tag{1.4}$$

Let A_n/B_n, A_n^1/B_n^1, and A_n^0/B_n^0 denote the n-th approximants of (1.1), (1.3), and (1.4), respectively. It follows that (Perron [1], p. 201)

$$\begin{aligned} A_n^0 &= A_{2n}, & A_n^1 &= A_{2n+1}, \\ B_n^0 &= B_{2n}, & B_n^1 &= B_{2n+1}, \end{aligned} \qquad (n = 0, 1, 2, \cdots).$$

Pringsheim has shown that the continued fraction (1.1)′ will converge if (Perron [1], p. 254)

$$|b_n| \geqq |a_n| + 1 \qquad (n = 1, 2, \cdots), \tag{1.5}$$

and that, subject to (1.5), the numbers $|B_n|$ associated with (1.1)′ satisfy the conditions

$$|B_n| - |B_{n-1}| \geqq |a_1a_2\cdots a_n| \qquad (n = 1, 2, 3, \cdots). \tag{1.51}$$

The plan of the proof of the theorem will be to show that if

$$\begin{aligned} &|1 + a_2| \geqq 1 + |a_1|, \\ &|a_{2n}| \leqq m < 1 && (n = 1, 2, 3, \cdots), \\ &|a_{2n+1}| \geqq \frac{2 + m}{1 - m} && (n = 1, 2, 3, \cdots), \\ &|a_{2n+1}| \geqq 2 + m + m\,|a_{2n-1}| && (n = 1, 2, 3, \cdots), \end{aligned} \tag{1.6}$$

conditions (1.5) are satisfied for (1.3) and (1.4). It will follow that both (1.3) and (1.4) converge. Thus the sequences $\{A_{2n+1}/B_{2n+1}\}$ and $\{A_{2n}/B_{2n}\}$ will converge. The proof will then be completed by proving that

$$\lim_{n\to\infty} \frac{A_{2n+1}}{B_{2n+1}} = \lim_{n\to\infty} \frac{A_{2n}}{B_{2n}}. \tag{1.7}$$

That the last two conditions (1.6) imply (1.5) follows at once from the observation that the conditions

$$|a_{2n+1}| - |1 + a_{2n}| \geqq 1 + |a_{2n-1}a_{2n}| \quad (n = 1, 2, 3, \cdots),$$

$$|a_{2n+1}| - |1 + a_{2n+2}| \geqq 1 + |a_{2n}a_{2n+1}| \quad (n = 1, 2, 3, \cdots),$$

together with the first condition (1.6), imply conditions (1.5) for both (1.3) and (1.4). The continued fractions (1.3) and (1.4) thus converge. It remains to demonstrate (1.7).

From (1.51) it follows that, subject to (1.6),

$$\begin{aligned} &|B_{2n+1}| - |B_{2n-1}| \geqq |a_1a_2 \cdots a_{2n}| \qquad (n = 1, 2, \cdots),\\ &|B_{2n+2}| - |B_{2n}| \geqq |a_1a_2 \cdots a_{2n+1}| \qquad (n = 0, 1, \cdots). \end{aligned} \tag{1.8}$$

Thus the sequences $\{|B_{2n}|\}$ and $\{|B_{2n+1}|\}$ are strictly increasing with n. Further, applying first the well-known formula for the difference between two consecutive approximants (Perron [1], p. 16) and then (1.8), we have

$$\left|\frac{A_{2n+1}}{B_{2n+1}} - \frac{A_{2n}}{B_{2n}}\right| = \left|\frac{a_1a_2 \cdots a_{2n+1}}{B_{2n}B_{2n+1}}\right| \leqq \frac{1}{|B_{2n+1}|}\left|\frac{B_{2n+2}}{B_{2n}} - 1\right|, \tag{1.9}$$

$$\left|\frac{A_{2n}}{B_{2n}} - \frac{A_{2n-1}}{B_{2n-1}}\right| = \left|\frac{a_1a_2 \cdots a_{2n}}{B_{2n-1}B_{2n}}\right| \leqq \frac{1}{|B_{2n}|}\left|\frac{B_{2n+1}}{B_{2n-1}} - 1\right|. \tag{1.10}$$

If the numbers $|B_{2n}|$ are uniformly bounded with respect to n, it follows that $\lim |B_{2n}|$ exists, $\neq 0$ and finite, and (1.7) follows at once from (1.9). If the numbers $|B_{2n+1}|$ are uniformly bounded, a similar argument shows that (1.7) follows from (1.10). There remains the possibility that $\lim |B_{2n}| = \lim |B_{2n+1}| = \infty$. To prove (1.7) for this case we recall that

$$B_{2n+2} = B_{2n+1} + a_{2n+2}B_{2n},$$

and hence that

$$\left|\frac{B_{2n+2}}{B_{2n}B_{2n+1}}\right| \leqq \frac{1}{|B_{2n}|} + \frac{|a_{2n+2}|}{|B_{2n+1}|} \leqq \frac{1}{|B_{2n}|} + \frac{1}{|B_{2n+1}|}.$$

Thus (1.7) follows from (1.9), and the proof of the theorem is complete.

In an entirely analogous manner one can show that *sufficient conditions for the convergence of* (1.1) *are the following*:

$$(1.11)\qquad |1 + a_2| \geqq |a_1| + 1, \qquad |a_2| \geqq \frac{2+m}{1-m},$$

$$|a_{2n+1}| \leqq m < 1, \qquad |a_{2n+2}| \geqq 2 + m + m\,|a_{2n}| \qquad (n = 1, 2, 3, \cdots),$$

where m is any positive number <1.

Conditions (1.6) and (1.11) are clearly independent of the sufficiency conditions referred to at the beginning of the paper, even when $m = \frac{1}{4}$. Further, they are not corollaries of (1.5) since conditions (1.5) never apply to continued fractions of the form (1.1). By means of (1.11) it is easy to construct examples to show that the continued fraction

$$1 + \frac{-\frac{1}{4} - \epsilon_1}{1+}\,\frac{-\frac{1}{4} - \epsilon_2}{1+}\,\frac{-\frac{1}{4} - \epsilon_3}{1+}\cdots$$

can converge for various choices of the $\epsilon_i > 0$, even though it diverges when $\epsilon_1 = \epsilon_2 = \epsilon_3 = \cdots = \epsilon > 0$.

In conclusion, we note that the following theorem can be established by a well-known argument (e.g., Perron [1], p. 260).

THEOREM. *If the quantities* a_n *are functions of a set of variables, the continued fraction* (1.1) *converges uniformly in any closed region characterized by the inequalities* (1.6) *or by* (1.11). *In particular, if the* a_n *are analytic functions of a single complex variable x, the continued fraction* (1.1) *represents a function f(x) which is analytic throughout the interior of the closed regions described.*

BIBLIOGRAPHY

W. LEIGHTON.

1. *A continued fraction expansion*, Bulletin of the American Mathematical Society, vol. 42(1936), p. 184, abstract.

W. LEIGHTON and H. S. WALL.

1. *On the transformation and convergence of continued fractions*, American Journal of Mathematics, vol. 58(1936), pp. 267–281.

O. PERRON.

1. *Die Lehre von den Kettenbrüchen*, Leipzig, 1929.

O. SZÁSZ.

1. *Über gewisse unendliche Kettenbruch-Determinanten und Kettenbrüche mit komplexen Elementen*, Sitzungsberichte der mathematisch-physikalischen Klasse der Akademie der Wissenschaften zu München (1912).
2. *Über die Erhaltung der Konvergenz Kettenbrüche bei independenter Veränderlichkeit aller ihre Elemente*, Journal für Mathematik, vol. 147(1916), pp. 132–160.

THE RICE INSTITUTE.

ALGEBRAIC FUNCTIONS OF ANALYTIC ALMOST PERIODIC FUNCTIONS

By Harald Bohr and Donald A. Flanders

It is our purpose here to extend to analytic almost periodic (a. p.) functions investigations previously carried on by several authors concerning the a. p. character of the solutions of algebraic equations whose coefficients are a. p. functions of a real variable.

The basic theorem (Theorem 1) concerning the existence of analytic a. p. solutions is an almost immediate consequence of the corresponding theorem (due to Walther and Cameron) for the real case, which latter may be stated as follows: If the coefficients of the equation

$$y^m + X_1(t)y^{m-1} + \cdots + X_m(t) = 0$$

are a. p. functions of the real variable t, and if the absolute value of the discriminant is bounded from zero, then the equation has m distinct a. p. solutions. Theorem 1 is proved in §1.

Further information concerning the nature of the solutions can be obtained by taking account of the exponents of the exponential series of the coefficients. Thus in the real case Cameron showed that the modulus of the Fourier exponents of each solution is contained in the quotient by an integer ($\leqq m$) of the modulus of the Fourier exponents of the coefficients. In a previous paper[1] we have refined this result by means of the notion of the "almost translation group" of the equation; particularly in the case where this group is transitive. As these results can be carried over directly to the analytic case, we shall not enter on them further, but refer the reader to the paper cited.

In the analytic case particular importance attaches to functions whose Dirichlet exponents are bounded on one side, since the functions are then a. p. in a half-plane. In §2 we show that when the exponents of the coefficient functions are bounded below and those of the discriminant have a minimum, then the exponents of the solutions are bounded below (Theorem 2). If in addition the exponents of the last coefficient likewise have a minimum, then the exponents of every solution also have a minimum.

Ostrowski[2] has treated the problem from quite another point of view, namely, the purely formal one, where the coefficients are taken to be formal Dirichlet

Received July 10, 1938.

[1] H. Bohr and D. A. Flanders, *Algebraic equations with almost-periodic coefficients*, Kgl. Danske Videnskabernes Selskab, Mathematisk-fysiske Meddelelser, vol. 15(1937), pp. 1–49.

[2] A. Ostrowski, *Über algebraische Funktionen von Dirichletschen Reihen*, Mathematische Zeitschrift, vol. 37(1933), pp. 98–133.

series without regard to the possibility of their representing functions. Such a formal treatment requires that the Dirichlet series considered be restricted to those of ordinary type, i.e., those with exponents bounded on one side and having no point of accumulation. The set of these series forms a field and by algebraic considerations Ostrowski succeeded in showing that this field is algebraically closed, that is, that every polynomial with coefficients in the field can be completely factored in the field. The formal and the analytic methods deal with the same material when the coefficients are a. p. functions whose Dirichlet series are of the ordinary type indicated above. After establishing in §3 some preliminary definitions and lemmas we devote §4 to showing that this class of functions, like the class of series considered by Ostrowski, forms an algebraically closed field (Theorem 3). There is appended a corollary, applying this theorem to the real variable case.

The setting of our Theorem 3 may be regarded as being derived from that of Ostrowski's problem by imposing upon the ordinary Dirichlet series appearing in the coefficients a restriction of function-theoretical nature, namely, that they shall be the Dirichlet series of a. p. functions. Ritt,[3] who was the first to extend the classical problem of algebraic functions of power series to exponential series, solved a similar problem by imposing a series-theoretical condition in that he supposed the coefficients to be ordinary Dirichlet series which are absolutely convergent in some half-plane. As before the set of series considered forms a field. In algebraic terminology the result of Ritt—which was independently found by Ostrowski and of which another proof is given in the paper cited above—states that this field, which is a subfield of the field dealt with in Theorem 3, is in its turn algebraically closed.

Both Ritt and Ostrowski consider other cases of algebraic functions of exponential series, which, however, do not connect immediately with our problem.

1. Let $s = \sigma + it$ be a complex variable. For $-\infty \leqq \alpha < \beta \leqq +\infty$ we shall denote by (α, β) the open strip of values of s given by $\alpha < \sigma < \beta$, $-\infty < t < +\infty$. For a similar choice of α and β we shall denote "every strip (α_1, β_1) such that $\alpha < \alpha_1 < \beta_1 < \beta$" by $[\alpha, \beta]$. Thus "$f(s)$ has property A in $[\alpha, \beta]$" means that $f(s)$ has this property in every strip (α_1, β_1), but not necessarily uniformly for all such strips. $[\alpha, \beta)$ and $(\alpha, \beta]$ have the corresponding meanings for $\alpha < \alpha_1 < \beta_1 \leqq \beta$ and $\alpha \leqq \alpha_1 < \beta_1 < \beta$, respectively.

We recall that a function $f(s)$ which is analytic in (α, β) and a. p. in $[\alpha, \beta]$ is bounded in $[\alpha, \beta]$; and that a rational integral function of such functions is itself analytic and a. p. in $[\alpha, \beta]$.

A function which is analytic and a. p. in $[\alpha, \beta]$ is an a. p. function of the real variable t on every vertical line $s = \sigma_0 + it$ in (α, β). The means to make the analytic case considered here depend upon the previously treated real case is a partial converse of this property, namely, a function which is analytic and bounded

[3] J. F. Ritt, *Algebraic combinations of exponentials*, Transactions of the American Mathematical Society, vol. 31 (1929), pp. 654–679.

in $[\alpha, \beta]$ and which is an a. p. function of t on *some* vertical line in (α, β) is a. p. in $[\alpha, \beta]$.

We shall use the fact that if $f(s)$ is analytic and a. p. in $[\alpha, \beta]$ and $\neq 0$ in (α, β), then GLB $|f(s)| > 0$ in $[\alpha, \beta]$.

Now let

$$y^m + x_1(s)y^{m-1} + \cdots + x_{m-1}(s)y + x_m(s) = 0 \tag{1}$$

be an algebraic equation of degree m in the complex variable y, with leading coefficient unity and remaining coefficients analytic a. p. functions of s in some common strip. We shall suppose that the discriminant

$$D(s) \equiv d[x_1(s), \cdots, x_m(s)] \tag{2}$$

is $\neq 0$ in some whole strip (α, β) lying in the foregoing common strip. Then equation (1) has a uniquely determined set of m distinct analytic solutions in (α, β), say

$$r_1(s), \cdots, r_m(s). \tag{3}$$

Since the leading coefficient of equation (1) is unity and the remaining coefficients are a. p. in $[\alpha, \beta]$ and hence bounded there, we know that each solution (3) is bounded in $[\alpha, \beta]$.

THEOREM 1. *If the coefficients of equation* (1) *are analytic in* (α, β) *and a. p. in* $[\alpha, \beta]$ *and* $D(s) \neq 0$ *in* (α, β), *then the m distinct analytic solutions* (3) *are a. p. in* $[\alpha, \beta]$.

As these solutions are bounded in $[\alpha, \beta]$ we have only to show that they are a. p. functions of t on some vertical line $s = \sigma_0 + it$ lying in (α, β). On such a line the coefficients of equation (1) are a. p. functions of t and the discriminant satisfies an inequality $|D(\sigma_0 + it)| > A > 0$, since $D(s)$ is a. p. $[\alpha, \beta]$ and $\neq 0$ in (α, β). Hence, applying the theorem for the real case mentioned in the introduction we see that the continuous solutions of this equation, i.e., the values for $s = \sigma_0 + it$ of the solutions (3), are a. p. functions of t.

2. The general result expressed in Theorem 1 becomes more available when we have some means of determining common strips of almost periodicity of the coefficients and zero-free strips of the discriminant. In an important class of cases information of this sort can be obtained from the Dirichlet series of the coefficients, to which we now turn. We shall denote the Dirichlet series of an analytic a. p. function $f(s)$ by $\sum_\lambda a_\lambda e^{\lambda s}$ (and write $f(s) \sim \sum_\lambda a_\lambda e^{\lambda s}$), where λ is a real parameter, a_λ is a one-valued complex function of λ which is zero for all but a countable set of values of λ, and those values of λ for which $a_\lambda \neq 0$ are the Dirichlet exponents of $f(s)$. The properties which we shall use may be summed up in the following two statements.

(a) If $f(s)$ is analytic in a half-plane $(-\infty, \beta)$, bounded in $(-\infty, \beta]$ and a. p. in $[-\infty, \beta]$, then it is a. p. in $(-\infty, \beta]$ and its Dirichlet exponents are non-nega-

tive, i.e., $f(s) \sim \sum_{\lambda \geqq 0} a_\lambda e^{\lambda s}$; and quasi-conversely, if $\sum_{\lambda \geqq 0} a_\lambda e^{\lambda(\alpha + it)}$ is the Fourier series of an a. p. function $F(t)$ of the real variable t, then there exists a function $f(s)$, analytic, a. p. and bounded in $(-\infty, \alpha)$, such that $f(s) \sim \sum_\lambda a_\lambda e^{\lambda s}$ and $f(\sigma + it) \to F(t)$ uniformly as $\sigma \to \alpha$.

(b) Let $f(s)$ be analytic, a. p. and bounded in $(-\infty, \beta]$ and let the GLB of its Dirichlet exponents be Λ ($\geqq 0$). There is a simple criterion by which we can determine whether Λ itself is an exponent (i.e., $a_\Lambda \neq 0$), namely: Λ is or is not an exponent according as there does or does not exist a half-plane $(-\infty, \gamma)$, throughout which $f(s) \neq 0$.

THEOREM 2. *If the Dirichlet exponents of the coefficients of equation* (1) *have a* GLB Λ, *while those of* $D(s)$ *have a minimum* M, *then the solutions* (3) *are a. p. in some* $[-\infty, \beta]$, *and their Dirichlet exponents are bounded below.*

For the moment consider the case where Λ (and hence M) is $\geqq 0$. Then the coefficients and $D(s)$ are analytic, a. p. and bounded in some $(-\infty, \gamma)$, and $D(s) \neq 0$ in some half-plane $(-\infty, \beta)$ lying in $(-\infty, \gamma)$. Hence the solutions are analytic and bounded in $(-\infty, \beta)$ and a. p. in $[-\infty, \beta]$ (hence in $(-\infty, \beta]$) so their Dirichlet exponents are non-negative. If now we let $\Lambda < 0$, we may transform equation (1) by multiplying its roots by $e^{-\Lambda s}$ into an equation of the sort just considered, whose roots will be analytic and a. p. in some $(-\infty, \beta]$, with their Dirichlet exponents non-negative. Then the product of each of these roots by $e^{\Lambda s}$ will be a root of the original equation, will have Λ as a lower bound of its exponents and will be a. p. at least in $[-\infty, \beta]$.

COROLLARY. *If further the Dirichlet exponents of the last coefficient* $x_m(s)$ *have a minimum, then so also do those of each solution in* $[-\infty, \beta]$.

For (as before multiplying the roots of the equation by $e^{-\Lambda s}$, if $\Lambda < 0$) if the exponents of some solution had no minimum, that solution would have zeros in every left half-plane. Then the product term, $x_m(s)$, would likewise have zeros in every half-plane, and hence could have no minimum exponent, contrary to hypothesis.

3. By a formal (generalized) Dirichlet series in s we shall mean a formal sum, $\sum_\lambda a_\lambda e^{\lambda s}$, where λ is a real parameter and a_λ (the coefficient function of the series) is a one-valued complex function of λ which is zero for all but a countable set of values of λ. Those values of λ for which a_λ is not zero will be called the exponents of the series. It will be convenient to denote such a series by a function symbol, say $a(s)$; we shall write $a(s) \approx \sum_\lambda a_\lambda e^{\lambda s}$ and say that $a(s)$ is formally equivalent to the series. Two formal series $a(s)$ and $b(s)$ are identical ($a(s) \equiv b(s)$) if and only if their coefficient functions are identical.

When a formal series $a(s)$ is the actual Dirichlet series of an a. p. function, we shall use the symbol $a(s)$ to denote the function as well.

If the exponents of $a(s)$, in their natural order as real numbers, form a discrete sequence, we shall attach indices 0, 1, 2, $\cdots$ to the successive values of λ in the

sequence, and shall denote a_{λ_n} by a_n. If the λ_n's have a finite point of accumulation Λ, we shall write $a(s) \approx \sum_{n=0}^{\lambda_n \uparrow \Lambda} a_n e^{\lambda_n s}$; if not, we shall write $a(s) \approx \sum_{n=0}^{\lambda_n \uparrow \infty} a_n e^{\lambda_n s}$. The latter type, which includes all Dirichlet polynomials, we shall call an ordinary Dirichlet series. The series whose coefficient function is identically zero, which we denote by 0, will be considered to be ordinary. In the notation for ordinary Dirichlet series it will often be convenient to allow $a_n = 0$ for some (or even all) values of n.

If $a(s) \approx \sum_\lambda a_\lambda e^{\lambda s}$, $b(s) \approx \sum_\lambda b_\lambda e^{\lambda s}$ are any two formal Dirichlet series, the one-valued function $c_\lambda = a_\lambda + b_\lambda$ is certainly zero for all but a countable set of values of λ. We shall call the series $c(s) \approx \sum_\lambda c_\lambda e^{\lambda s}$ the sum of $a(s)$ and $b(s)$, and write $c(s) \equiv a(s) + b(s)$.

If

$$a(s) \approx \sum_{n=0}^{\lambda_n \uparrow \infty} a_n e^{\lambda_n s} \quad \text{and} \quad b(s) \approx \sum_{n=0}^{\mu_n \uparrow \infty} b_n e^{\mu_n s}$$

are any formal ordinary Dirichlet series, their product $d(s) \equiv a(s)\, b(s)$ is defined to be the series $\sum_\nu d_\nu e^{\nu s}$, where d_ν is the sum of the (certainly finite) set of products $a_\lambda b_\mu$ for which $\lambda + \mu = \nu$. The series $d(s)$ is also ordinary and it is easily seen that the set of all formal ordinary Dirichlet series forms a field.

Ostrowski's result for formal ordinary Dirichlet series may be put in the following form (negligibly modified for our purposes).

Ostrowski's Theorem. *If*

$$y^m + x_1(s) y^{m-1} + \cdots + x_m(s) \tag{4}$$

is a polynomial in y with leading coefficient unity and remaining coefficients formal ordinary Dirichlet series in s,

$$x_h(s) \approx \sum_{n=0}^{\lambda_n \uparrow \infty} a_{h,n} e^{\lambda_n s} \qquad (h = 1, \cdots, m),$$

then there exists a unique set of m formal ordinary Dirichlet series,

$$y_j(s) \approx \sum_{n=0}^{\mu_n \uparrow \infty} b_{j,n} e^{\mu_n s} \qquad (j = 1, \cdots, m),$$

such that

$$y^m + x_1(s) y^{m-1} + \cdots + x_m(s) \equiv \prod_{j=1}^{m} [y - y_j(s)], \tag{5}$$

i.e., such that

$$x_h(s) \equiv (-1)^h \sum_{j_1 < \cdots < j_h} y_{j_1}(s) \cdot \cdots \cdot y_{j_h}(s) \quad (h = 1, \cdots, m). \tag{6}$$

Since the ordinary Dirichlet series form a field, relation (5) is true when any such formal series $r(s)$ is substituted for y; and the vanishing of $\prod_{j=1}^{m} [r(s) - y_j(s)]$ implies the vanishing of one of the factors. Thus the polynomial (4) has just m (not necessarily distinct) zeros in the field of formal ordinary Dirichlet series.

As is well known the Dirichlet series of the product $d(s)$ of two analytic a. p. functions in a common strip, say $a(s) \sim \sum_\lambda a_\lambda e^{\lambda s}$, $b(s) \sim \sum_\mu b_\mu e^{\mu s}$, is the series $\sum_\nu d_\nu e^{\nu s}$ which is derived from the series $a(s)$ and $b(s)$ by the rule just given for ordinary Dirichlet series; in this case $\sum_{\lambda+\mu=\nu} a_\lambda b_\mu$ may be an infinite series which however is always unconditionally convergent. While the set of all Dirichlet series of analytic a. p. functions in a common strip form a ring merely, not a field, still the set of all formal ordinary Dirichlet series which are the actual Dirichlet series of analytic a. p. functions (each in some half-plane) does form a field.

For the class of all formal Dirichlet series the above definition of multiplication fails to give a result in general, and even where it applies, this multiplication does not obey the usual rules. For example, the product of $a(s) \approx 1 - e^s$ and $b(s) \approx \sum_{n=-\infty}^{+\infty} e^{ns}$ is 0, although neither factor is 0; and the product of $a(s)b(s)$ by $b(s)$ is 0 but that of $a(s)$ by $[b(s)]^2$ does not exist since $[b(s)]^2$ fails to exist. While our Theorem 3 deals only with elements in the field of formal ordinary Dirichlet series, we shall need in the proof to operate with generalized Dirichlet series of other types than those for which multiplication has been defined above. For our purposes, however, the following limited definition of the product of m (>1) formal Dirichlet series suffices, since it gives the two succeeding lemmas, which are all that we shall need.

In the case of two series $a_1(s) \approx \sum_{\lambda_1} a_{1,\lambda_1} e^{\lambda_1 s}$ and $a_2(s) \approx \sum_{\lambda_2} a_{2,\lambda_2} e^{\lambda_2 s}$ we follow the definition above, i.e., we say that the product $a_1(s)a_2(s)$ exists and is equal to $d(s) \approx \sum d_\lambda e^{\lambda s}$, if for any fixed λ the sum $d_\lambda = \sum_{\lambda_1+\lambda_2=\lambda} a_{1,\lambda_1} a_{2,\lambda_2}$ is a finite or an unconditionally convergent infinite series. Generally, if $a_h(s) \approx \sum a_{h,\lambda_h} e^{\lambda_h s}$ $(h = 1, \cdots, m)$ are m formal Dirichlet series, we define the product $a_1(s) \cdots a_m(s)$ as the repeated product

$$(\cdots((a_1(s)a_2(s))a_3(s)) \cdots)a_m(s)$$

when it exists. This amounts to defining $a_1(s) \cdots a_m(s)$ by $\sum d_\lambda e^{\lambda s}$, whenever the coefficient function

$$d_\lambda = \sum_{\lambda_m} (\cdots (\sum_{\lambda_3} (\sum_{\lambda_2} a_{1,\lambda-\lambda_m-\cdots-\lambda_2} \cdot a_{2,\lambda_2}) a_{3,\lambda_3}) \cdots) a_{m,\lambda_m}$$

exists for all λ.

LEMMA 1. *To Ostrowski's theorem add the assumption that the coefficients $x_\lambda(s)$ are the actual Dirichlet series of analytic a. p. functions. Let*

$$r(s) \sim \sum_{\nu \geqq N} r_\nu e^{\nu s}$$

be an analytic a. p. function, whose Dirichlet exponents are bounded below by N. *If the series* $r(s)$ *be substituted for* y *in relation* (5), *then the formal series expressed by each side of the relation exist, and the relation*

$$(7)\qquad [r(s)]^m + x_1(s)[r(s)]^{m-1} + \cdots + x_m(s) \equiv \prod_{j=1}^{m} [r(s) - y_j(s)]$$

holds.

Proof. As is easily seen, no generality will be lost if we suppose $\lambda_0 \geqq 0$, $\mu_0 \geqq 0$ and $\mathrm{N} \geqq 0$, so that all exponents involved are non-negative.

The formal series on the left exists since all the series involved are the actual Dirichlet series of analytic a. p. functions in a common half-plane. Let c_ξ denote its coefficient function.

Set

$$r(s) - y_j(s) \approx \sum_{\xi_j} g_{j,\xi_j} e^{\xi_j s} \qquad (j = 1, \cdots, m).$$

We shall prove that the coefficient function d_ξ of the product of these series exists and is identical with c_ξ.

As the exponents in each factor $r(s) - y_j(s)$ are $\geqq 0$, for every fixed ξ all coefficients g_{j,ξ_j} involved in the determination of d_ξ will have their indices $\xi_j \leqq \xi$, so that the determination of d_ξ will not be affected if we alter any of the terms with exponent $> \xi$ in the factors $r(s) - y_j(s)$. Since $\mu_n \uparrow \infty$, there exists an index p such that $\mu_n > \xi$ for $n > p$. If then in each factor $r(s) - y_j(s)$ we replace $y_j(s)$ by

$$y_j^*(s) \approx \sum_{n=0}^{p} b_{j,n} e^{\mu_n s},$$

the determination of the value for ξ of the coefficient function of

$$(8)\qquad \prod_{j=1}^{m} [r(s) - y_j^*(s)]$$

will be identical with the determination of d_ξ. Since each $y_j^*(s)$, being a Dirichlet polynomial, is the Dirichlet series of an analytic a. p. function in $(-\infty, +\infty]$, each factor $r(s) - y_j^*(s)$ is the Dirichlet series of an analytic a. p. function in some half-plane, so d_ξ exists. We now proceed to prove that $d_\xi = c_\xi$. If we set

$$x_h^*(s) = (-1)^h \sum_{j_1 < \cdots < j_h} y_{j_1}^*(s) \cdots y_{j_h}^*(s) \qquad (h = 1, \cdots, m),$$

the series (8) is identical with the series

$$(9)\qquad [r(s)]^m + x_1^*(s)[r(s)]^{m-1} + \cdots + x_m^*(s)$$

since both represent the same a. p. function. But for our fixed ξ the coefficient d_ξ in (9) is in turn equal to the corresponding coefficient c_ξ in

$$[r(s)]^m + x_1(s)[r(s)]^{m-1} + \cdots + x_m(s),$$

since the polynomials $x_h^*(s)$ can differ from the series $x_h(s)$ only in terms with exponents $> \xi$.

LEMMA 2. *Let $a_1(s), \cdots, a_m(s)$ be m (>1) formal Dirichlet series each of which either is identically zero or has a minimum exponent. If the product $a_1(s) \cdots a_m(s)$ exists and is 0, one of the factors must be 0.*

Proof. If no factor were 0, the product of the terms with minimum exponents would be a non-vanishing term in the product $a_1(s) \cdots a_m(s)$.

4. As already mentioned, the set of all formal ordinary Dirichlet series, each of which is the actual Dirichlet series of an analytic a. p. function in some left half-plane, forms a subfield $\mathfrak{F}$ of Ostrowski's field of all formal ordinary Dirichlet series. In order to show that the field $\mathfrak{F}$ (like Ostrowski's field) is algebraically closed, i.e., that every polynomial over $\mathfrak{F}$ is completely factorable in $\mathfrak{F}$, we need concern ourselves only with those polynomials whose discriminants do not vanish identically. For a polynomial over a given field with null discriminant is factorable in that field. Expressed in terms of the solutions of equations, we must prove the following theorem.

THEOREM 3. *If the coefficients of equation* (1) *are a. p. functions in a half-plane with Dirichlet series of the ordinary type and with not identically vanishing discriminant, then the analytic solutions* (3), *a. p. in some* $[-\infty, \beta]$, *have likewise ordinary Dirichlet series.*

Since our field $\mathfrak{F}$ is a subfield of Ostrowski's field, equation (1) has a unique set of formal solutions in the latter field, and we have to show that these formal series are the actual Dirichlet series of the analytic a. p. solutions in $[-\infty, \beta]$. Also, it is clearly sufficient to prove that every equation given by our hypothesis has at least one root in $\mathfrak{F}$, since the "reduced" equation will still have coefficients in $\mathfrak{F}$ with discriminant not identically zero. Hence we shall prove the theorem indirectly by showing that the assumption that no root of equation (1) lies in $\mathfrak{F}$ leads to a contradiction.

Let $r(s)$ be one of the (a. p.) roots of the equation (1). Since the last coefficient $x_m(s)$ and the discriminant $D(s)$ have minimum exponents this will likewise be the case with $r(s)$ so that we may write

$$r(s) \sim c_0 e^{\nu_0 s} + \sum_{\nu > \nu_0} c_\nu e^{\nu s}.$$

We now transform equation (1) by diminishing its roots by $c_0 e^{\nu_0 s}$. Since this involves only rational integral operations on this term and the original coefficients, neither the domain of almost periodicity of the coefficients nor the character of their Dirichlet series will be changed; furthermore, the last coefficient can not be identically zero (as none of the roots are polynomials). Hence the Dirichlet series of the solution $r(s) - c_0 e^{\nu_0 s}$ of the new equation will have a minimum exponent, so that

$$r(s) \sim c_0 e^{\nu_0 s} + c_1 e^{\nu_1 s} + \sum_{\nu > \nu_1} c_\nu e^{\nu s}.$$

If we proceed in this way, it is clear that our solution $r(s)$ has a series of the form

$$r(s) \sim \sum_{n=0}^{\nu_n \uparrow N^*} c_n e^{\nu_n s} + \sum_{\nu \geqq N^*} c_\nu e^{\nu s}, \tag{10}$$

where the first summation contains infinitely many non-zero terms. If now we apply Lemma 1 of §3, taking for the $r(s)$ entering there the solution (10) of equation (1), we get

$$\prod_{j=1}^{m} [r(s) - y_j(s)] \equiv [r(s)]^m + x_1(s)[r(s)]^{m-1} + \cdots + x_m(s) \equiv 0.$$

But since the series $y_j(s)$ are ordinary, only a finite number of their terms have exponents $< N^*$. Hence each factor $r(s) - y_j(s)$ must have a minimum exponent. By Lemma 2, this requires that some $r(s) - y_j(s) \equiv 0$. This contradicts our assumption that the Dirichlet series of $r(s)$ was not ordinary.

This theorem has a corollary in the real case which does not seem to be easy to obtain without using the theory of analytic functions:

COROLLARY. *Let*

$$y^m + X_1(t)y^{m-1} + \cdots + X_m(t) = 0 \tag{11}$$

be an algebraic equation whose coefficients $X_h(t)$ are a. p. functions of the real variable t with ordinary Fourier series (i.e., $\lambda_n \uparrow \infty$) and let its discriminant be not identically zero. Then if the equation has as a solution an a. p. function $R(t)$ whose Fourier exponents are bounded below, the Fourier series of this solution is ordinary.

According to a well known theorem quoted above each $X_h(t)$ is the boundary value for $s = 0 + it$ of an analytic a. p. function $x_h(s)$ in $[-\infty, 0)$ whose exponents and coefficients are identical with those of $X_h(t)$; and similarly there exists an analytic function $r(s)$ a. p. in $[-\infty, 0)$ whose boundary value is the solution $R(t)$ in question. Hence (a) the equation

$$y^m + x_1(s)y^{m-1} + \cdots + x_m(s) = 0 \tag{12}$$

has analytic a. p. solutions in some $[-\infty, \beta]$ with ordinary Dirichlet series and (b) since $R(t)$ satisfies (12) for $\sigma = 0$, its series formally satisfies (12) for $\sigma = 0$. Thus the series of $r(s)$ formally satisfies (12), so $r(s)$ is a solution of (12) in $[-\infty, \beta]$. From Theorem 3 we conclude that the Dirichlet series of $r(s)$ is ordinary; hence the Fourier series of $R(t)$, which has the same exponents, is likewise ordinary.

UNIVERSITY OF COPENHAGEN.

DISCONTINUOUS GROUPS AND ALLIED TOPICS, III: ON A LEMMA ABOUT MATRICES

By Max Zorn

1. **Introduction**. In connection with the theory of hypercomplex units the following problem arises. Given a linear family $\mathfrak{A}$ of *real* matrices, which contains with every A its transpose A^T and with every regular A its inverse A^{-1}. Supposing that $\mathfrak{A}$ contains at least one unimodular matrix A_1, $|A_1| = 1$, to determine the unimodular matrices in $\mathfrak{A}$ for which the sum of the squares of the elements is a minimum. The results and proofs are extremely simple. The minimizing matrices are identical with the orthogonal matrices in the family. This is shown by application of Lagrange's multiplier rule. The more or less algebraic machinery may very well be extended to more general linear spaces. A. D. Michal expresses the opinion that the full theory might be generalized; yet it seems that at present the necessary existence theorems are not available. Apart from possible arithmetical applications the existence of at least one orthogonal matrix in such families is of a certain independent interest.

2. **Notations**. The determinant $|a_{ik}|$ of the matrix $A = (a_{ik})$ is as usual denoted by $|A|$, the inverse by A^{-1}, the transpose (a_{ki}) by A^T. The indices i, k range from 1 to n, the number of rows or columns of the matrices. The Greek index ranges from 1 to m, unless otherwise stated. The trace $t(A)$ is the sum $\sum a_{ii}$ of the diagonal elements. For example, the trace $t(I)$ of the unit matrix is equal to n.

A linear combination $\sum_{i,k} c_{ik}x_{ik}$ of symbols x_{ik} with coefficients c_{ik} is expressible in the form $t(XC^T)$, if we let $X = (x_{ik})$, $C = (c_{ik})$.

The sum of the squares of all elements is $t(XX^T)$. Since we shall deal only with real matrices, $t(XX^T)$ is positive and vanishes only for $X = 0$. A matrix A is orthogonal if $X^{-1} = X^T$ and if $|X| = 1$. For an orthogonal matrix XX^T is I, and $t(XX^T) = n$.

3. **Differentials**. We shall use x_{ik} as independent variables and dx_{ik} as their differentials.

For a matrix $F = (f_{ik})$ of functions we define $dF = (df_{ik})$. In particular we shall have to make use of

$$dX = d(x_{ik}) = (dx_{ik}),$$

$$dt(XC^T) = t(dXC^T),$$

$$dt(XX^T) = 2t(X^TdX).$$

Received August 17, 1938; presented to the American Mathematical Society, April 9, 1938; cf. Bulletin of the American Mathematical Society, vol. 44(1938), page 354.

Apart from these trivial formulas we need the differential of the determinant $| X |$:

$$d| X | = \sum_{i,k} \frac{\partial | X |}{\partial x_{ik}} dx_{ik} = t\left(dx\left(\frac{\partial | X |}{\partial x_{ik}}\right)^T\right)$$

$$= t(dX \cdot X^{-1} | X |) = | X | \, t(dX \cdot X^{-1}).$$

In the last two expressions values may be substituted for the x_{ik} provided that the value of $| X |$ is not zero. We are here mainly concerned with the case $| X | = 1$, where

$$d \,| X | = t(dX \cdot X^{-1}).$$

4. **Linear families.** We consider linear families $\mathfrak{A}$ of real matrices over the field of real numbers. That is, we have a set $\mathfrak{A}$ of matrices A, B such that

(I) the coefficients a_{ik}, b_{ik} are real and for arbitrary real numbers a, b $aA + bB$ is in the set.

These sets are further restricted by the properties:

(II) If A is in $\mathfrak{A}$, then the transpose A^T is also in $\mathfrak{A}$.

(III) If A is in $\mathfrak{A}$ and $| A | \neq 0$ (or if A^{-1} exists), then A^{-1} is in $\mathfrak{A}$.

(IV) There exists one unimodular matrix A_1 in $\mathfrak{A}$, $| A_1 | = 1$.

Without loss of generality we may assume that $\mathfrak{A}$ is defined by a set of independent *linear* relations for its elements

$$\sum_{i,k} c_{ik\nu} x_{ik} = 0 \qquad (\nu = 1, \cdots, m).$$

These we write, using the matrices

$$C_\nu = (c_{ik\nu}) \qquad (\nu = 1, \cdots, m),$$

in the form

$$t(C_\nu^T X) = 0 \qquad (\nu = 1, \cdots, m).$$

5. **The minimizing matrices.** The Lagrange multiplier rule establishes a necessary condition for a (relative) minimum of a function $f(x_{ik})$ of real variables x_{ik} at a "point" a_{ik} with the side conditions

$$\varphi_0(x_{ik}) = c_0, \qquad \varphi_\nu(x_{ik}) = c_\nu \qquad (\nu = 1, \cdots, m).$$

(Indices are chosen such that they can be used immediately for our special problem.)

If the rank of the matrix

$$\left(\frac{\partial \varphi_\nu}{\partial x_{ik}}\right) \qquad (\nu = 0, 1, \cdots, m)$$

is $m + 1$, or in other words if the differentials $d\varphi_0$, $d\varphi_\nu$ are linearly independent at a_{ik}, then a necessary condition for a minimum is the existence of constants λ_0, λ_ν such that

$$df + \sum_{\nu=1}^{m} \lambda_\nu d\varphi_\nu = 0$$

for $x_{ik} = a_{ik}$.

In the present case we want

$$\varphi_0(x_{ik}) = | x_{ik} | = | X | = 1,$$
$$\varphi_\nu(x_{ik}) = t(C_\nu^T X) = 0.$$

The function f is

$$f(x_{ik}) = \sum x_{ik}^2 = t(XX^T).$$

For λ_0 we shall also write $-\lambda$. For a matrix A which minimizes $t(XX^T)$ we obtain consequently (postponing the independence proof)

$$d(XX^T) + \lambda_0 d | X | + \sum \lambda_\nu dt(C_\nu^T X) = 0 \qquad (\text{for } X = A).$$

Using the relations in §3, we have

$$2t(X^T dX) + \lambda_0 | X | t(X^{-1} dX) + t(\sum \lambda_\nu C_\nu^T dX) = 0$$

for $X = A$, provided that $| A | \neq 0$; and for $| A | = 1$ we get

$$t(\{2A^T + \lambda_0 A^{-1} + \sum \lambda_\nu C_\nu^T\} dX) = 0.$$

Since the dx_{ik} are independent variables, the last formula implies that

$$2A^T + \lambda_0 A^{-1} + \sum \lambda_\nu C_\nu^T = 0.$$

Now consider the matrix

$$D = 2A^T + A^{-1}.$$

We know that $D = \sum - \lambda_\nu C_\nu^T$. From these two equations it follows that D is equal to zero. First of all, D is in the family, for it is a linear combination of A^T and A^{-1}, which are both members of $\mathfrak{A}$ because of properties (II) and (III). Property (II) implies then that D^T is in $\mathfrak{A}$. $\mathfrak{A}$ was defined by the linear relations $t(C_\nu^T X) = 0$, whence

$$t(C_\nu^T D^T) = 0.$$

Finally we utilize the second equality:

$$t(D^T D) = t(D^T \sum - \lambda_\nu C_\nu^T) = -\sum t(D^T C_\nu^T) = 0.$$

Since D is a real matrix, $t(DD^T) = 0$ guarantees $D = 0$, and we have derived the equation

$$2A^T + \lambda_0 A^{-1} = 0.$$

The orthogonality of A now follows immediately. We have

$$2A^T = +\lambda A^{-1}, \qquad 2AA^T = \lambda I,$$

$0 \leqq t(2AA^T) = \lambda n$, hence $\lambda \geqq 0$. If we consider the determinants, we get

$$2^n \mid A^T \mid = 2^n = \lambda^n \mid A^{-1} \mid = \lambda^n;$$

from $\lambda \geqq 0$ and $\lambda^n = 2^n$ it follows that $\lambda = 2$, and

$$A^T = A^{-1}.$$

Since $\mid A \mid$ was supposed to be $= 1$, A is orthogonal.

We have now to show that the $d\varphi$ are linearly independent. This is done with the same apparatus. A linear relation with the constants μ_0, μ_ν :

$$\mu_0 d \mid X \mid + \sum \mu_\nu \, dt(C_\nu^T X) = 0,$$

for $X = A$, $\mid A \mid = 1$, implies

$$t(\{\mu_0 A^{-1} + \sum \mu_\nu C_\nu^T \, dX\}) = 0, \qquad \mu_0 A^{-1} = \sum - \mu_\nu C_\nu^T.$$

It follows that

$$t(\mu_0 A^{-1} A^{-T}) = \sum - \mu_\nu t(A^{-T} C_\nu^T) = 0, \qquad \mu_0 t(A^{-1} A^{-T}) = 0,$$

and from $A^{-1} \neq 0$, $\mu_0 = 0$. A relationship $0 = \sum - \mu_\nu C_\nu^T$ is only possible for $\mu_\nu = 0$, since the C_ν^T are, like the C_ν, linearly independent.

We notice that the extremal property of A has not been used; throughout the set $\mathfrak{M}$ of unimodular matrices in $\mathfrak{A}$ the differentials of the defining functions are independent. The theory of implicit functions informs us that under these circumstances $\mathfrak{M}$ is locally Euclidean. It may consist of isolated points.

6. **Existence theorem.** The function $t(XX^T)$ may be interpreted as a squared distance of X from the zero matrix in an n^2-dimensional Euclidean space. The unimodular matrices of a linear family $\mathfrak{A}$ form a closed subset $\mathfrak{M}$; consequently, there exists at least one matrix A of (absolutely) minimal $t(XX^T)$; of course $t(AA^T)$ is $\leqq t(A_1 A_1^T)$. In view of the foregoing proof we may state the theorem:

Every family of (finite, square) matrices with properties (I), (II), (III), (IV) *contains an orthogonal matrix.*

Choosing $\mathfrak{A}$ as the set of all matrices, we have:

If $\mid A \mid = 1$, *then* $t(AA^T) \geqq n$; *the equality holds exactly in the case of orthogonal matrices.*

Since $\mathfrak{M}$ is locally Euclidean, it is locally connected and the components $\mathfrak{N}$ of $\mathfrak{M}$ (the largest connected subsets) are open. Since a component is closed, it contains an element which yields within $\mathfrak{N}$ a minimum for $t(XX^T)$; and since $\mathfrak{N}$ is open in $\mathfrak{M}$, this is a relative minimum with respect to $\mathfrak{M}$. Therefore we have:

Every component of $\mathfrak{M}$ contains at least one orthogonal matrix.

Let $\{\mathfrak{N}_i\}$ be the set of all components; then $\mathfrak{N}_i$ contains an orthogonal matrix

A_i ; and since the set of all orthogonal matrices is compact (i.e., every infinite set of orthogonal matrices contains a sequence which converges towards an orthogonal matrix), we conclude:

$\mathfrak{M}$ *has a finite number of components.*

For if we take a convergent sequence of orthogonal matrices converging to an orthogonal matrix A_ω, almost all of them must be contained in the component of A_ω so that only a finite number of A_i from different components is possible.

University of California at Los Angeles.

THE SUMMABILITY OF EXPONENTIAL AND FACTORIAL SERIES

By Tomlinson Fort

The summability of ordinary factorial series by Cesàro means has been studied by Bohr.[1] However, nothing seems to have been done since the appearance of Bohr's paper. In §§2 and 3 of the present paper a study is made of what is called R-summability of general factorial series and a type of related exponential series. A general theorem on R-summability is proved in §1. This is one of the most interesting theorems of the paper.

1. ***R*-summability**. Let there be given a sequence $0 < \lambda_n \to \infty$. Let

$$\sigma_n = \sum_{n=1}^{n} \frac{1}{\lambda_n}$$

and let $\sigma_n \to \infty$. Also let there be given a series

$$\sum_{n=1}^{\infty} a_n(z). \tag{1}$$

Let

$$S_n^{(0)}(z) = \sum_{n=1}^{n} a_n(z)$$

and

$$S_n^{(k)}(z) = \frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n} S_n^{(k-1)}(z), \qquad n > 0.$$

Definition.[2] *We call $S_n^k(z)$ the k-th R-mean for series* (1). *If* $\lim_{n\to\infty} S_n^{(k)}(z)$ *exists and equals $S(z)$, we say that* (1) *is summable $R[k, \lambda]$ to $S(z)$. Summation is said to be uniform over a set P in case $S_n^{(k)}(z)$ approaches its limit uniformly over P.*

We note without proof the following three readily established theorems with reference to R-summability.

Theorem A. *If a series is uniformly summable $R[k-1, \lambda]$ over P to $s(z)$, then it is uniformly summable $R[k, \lambda]$ over P to $s(z)$.*

Received October 3, 1938.

[1] Gött. Nach., 1909, p. 260.

[2] M. Riesz (Comptes Rendus, vol. 149, p. 18) introduces weighted means as a generalization of summability by the method of the arithmetic mean of the first order. He generalizes to his "typical means" of arbitrary order. R-summability as defined here generalizes Riesz means of the first order by simple iteration as the Hölder method generalizes the ordinary arithmetic mean of the first order.

Theorem B. *If* (1) *is summable* $R[k, \lambda]$ *to* s, *then* $\sum_{n=m}^{\infty} a_n$ *is summable* $R[k, \lambda]$ *to* $s - \sum_{n=1}^{m-1} a_n$, *and conversely.*

Theorem C. *If* $a_n > 0$ *and* (1) *is summable* $R[k, \lambda]$, *then* (1) *is convergent.*

Consider now a series of the form

$$\sum_{n=1}^{\infty} a_n(z) b_n(z). \tag{2}$$

Denote the successive R-means formed for (1) by $s_n^k(z)$ and for (2) by $S_n^{(k)}(z)$. We proceed, successively summing by parts, that is, applying the following general formula

$$\sum_{n=0}^{n} u_n w_n = u_{n+1} \sum_{n=0}^{n} w_n - \sum_{n=0}^{n} (\Delta u_n) \sum_{n=0}^{n} w_n. \tag{3}$$

First we obtain

$$S_n^{(0)}(z) = b_{n+1}(z) s_n^{(0)}(z) - \sum_{n=1}^{n} (\Delta b_n(z)) s_n^{(0)}(z). \tag{4}$$

The operator Δ here, as throughout the paper, applies to the variable n. Apply the operator

$$\frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n}$$

to both members of (4) and sum by parts in each instance, letting

$$\frac{1}{\lambda_n} s_n^{(0)}(z) = w_n$$

of formula (3). We get

$$\begin{aligned} S_n^{(1)}(z) = b_{n+2}(z) s_n^{(1)}(z) &- \frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n} [\lambda_{n+1} \Delta b_n(z)] \sigma_n s_n^{(1)}(z) \\ &- \frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n} [\lambda_{n+1} \Delta b_{n+1}(z)] \sigma_n s_n^{(1)}(z) \\ &+ \frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n} \sum_{n=1}^{n} [\Delta \lambda_n \Delta b_n(z)] \sigma_n s_n^{(1)}(z). \end{aligned} \tag{5}$$

From this formula one can conclude the following theorem.

Theorem I. *If* (1) *is uniformly summable* $R[1, \lambda]$ *over a set* P, *if* $b_n(z)$ *and* $(\lambda_n \Delta b_n(z))\sigma_n$ *each approaches a limit uniformly over* P, *if* $(\lambda_{n+1}/\lambda_n) \to C$ *and if* $\sum_{n=1}^{\infty} | \Delta\lambda_n \Delta b_n(z) | \sigma_n$ *converges uniformly over* P, *then* (2) *is uniformly summable* $R[1, \lambda]$ *over* P.

We proceed to extend this theorem. Apply the operator

$$\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}$$

to both members of (5). Sum by parts exactly as we did to obtain (5) from (4). We obtain

$$
\begin{aligned}
S_n^{(2)}(z) &= b_{n+3}(z)s_n^{(2)}(z) - \frac{1}{\sigma_n}\sum_{n=1}^{n}[\Delta b_{n+2}(z)]\sigma_n s_n^{(2)}(z) \\
&\quad - \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{\lambda_{n+1}}{\lambda_n}[\Delta b_n(z)]\sigma_{n+1}s_n^{(2)}(z) \\
&\quad - \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{\lambda_{n+2}}{\lambda_n}[\Delta b_{n+2}(z)]\sigma_{n+1}s_n^{(2)}(z) \\
&\quad + \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\frac{1}{\sigma_n}\sum_{n=1}^{n}[\Delta(\lambda_n(\Delta b_n(z))\sigma_n)]\sigma_n s_n^{(2)}(z) \qquad (6)\\
&\quad + \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\frac{1}{\sigma_n}\sum_{n=1}^{n}[\Delta(\lambda_{n+1}(\Delta b_{n+1}(z))\sigma_{n+1})]\sigma_n s_n^{(2)}(z) \\
&\quad + \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{\lambda_{n+1}}{\lambda_n}[\Delta(\lambda_{n+1}(\Delta b_{n+1}(z))\sigma_{n+1})]\sigma_n s_n^{(2)}(z) \\
&\quad - \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\sum_{n=1}^{n}[\Delta(\lambda_n(\Delta\lambda_n\Delta b_n(z))\sigma_n)]\sigma_n s_n^{(2)}(z).
\end{aligned}
$$

From (6) we can deduce the following theorem.

THEOREM II. *If*

(i) *series* (1) *is uniformly summable* $R[2, \lambda]$ *over* P,

(ii) $\sum_{n=1}^{\infty} | \Delta(\lambda_n(\Delta\lambda_n\Delta b_n(z))\sigma_n) | \sigma_n$ *converges uniformly over* P,

(iii) $b_n(z)$, $\lambda_n(\Delta b_n(z))\sigma_n$, $\lambda_n[\Delta(\lambda_n(\Delta b_n(z))\sigma_n)]\sigma_n$ *all approach limits uniformly over* P, *and if*

(iv) $(\lambda_{n+1}/\lambda_n) \to C$,

then (2) *is uniformly summable over* P.

Complete generalization of Theorem I is now easy. Apply the operator

$$\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}$$

to both sides of (6) and sum by parts exactly as we have previously done. Repeat until the left member is the r-th mean for (2). We do not write out the resulting expression. However, we state the following theorem.

Let $L_n^{(0)} = b_n(z)$ and let $L_n^{(k)} = \lambda_n\sigma_n\Delta L_n^{(k-1)}$ when $k > 0$. Also let $M_n^{(0)} = \Delta b_n(z)$ and let $M_n^{(k)} = \sigma_n\Delta\lambda_n M_n^{(k-1)}$ when $k > 0$.

THEOREM III. *If*

(i) *series* (2) *is uniformly summable* $R[r, \lambda]$ *over* P,

(ii) $L_n^{(k)}$ $(k = 1, \cdots, r)$ *approaches a limit uniformly over* P,

(iii) $\sum_{n=1}^{\infty} | M_n^{(r)} |$ *converges uniformly over* P,

(iv) $(\lambda_{n+1}/\lambda_n) \to C$,

then series (2) *is uniformly summable* $R[r, \lambda]$ *over* P.

2. **Exponential series.** Consider the exponential series

$$\sum_{n=1}^{\infty} C_n e^{p_{1n}}, \tag{7}$$

where

$$p_{1n} = -z \sum_{n=1}^{n} \frac{1}{\lambda_n} + \frac{z^2}{2} \sum_{n=1}^{n} \frac{1}{\lambda_n^2} - \cdots + (-1)^w \frac{z^w}{w} \sum_{n=1}^{n} \frac{1}{\lambda_n^w},$$

where $0 < \lambda_n \to \infty$ and $\sum_{n=1}^{\infty} \frac{1}{\lambda_n}$ is divergent and

$$\frac{\lambda_{n+1}}{\lambda_n} = 1 + \frac{e_1}{\lambda_n} + \frac{e_2}{\lambda_n^2} + \cdots + \frac{e_r}{\lambda_n^r} + \frac{e_{r+1}(n)}{\lambda_n^{r+1}},$$

where $e_1, \cdots, e_r$ are constants and e_{r+1} is bounded. We shall prove the following theorem.

THEOREM IV. *If* (7) *is summable* $R[r, \lambda]$ *at* z_0, *then it is summable* $R[r, \lambda]$ *at each point* z *which is so situated that* Real $(z) >$ Real (z_0) *and uniformly over any bounded region* P *at all points of which* Real $(z) >$ Real $(z_0) + \epsilon$.

Proof is by means of Theorem III. We let

$$L_n^{(0)}(z) = b_n(z) = e^{p_{2n}}, \tag{8}$$

where

$$p_{2n} = -(z - z_0) \sum_{n=1}^{n} \frac{1}{\lambda_n} + \frac{z^2 - z_0^2}{2} \sum_{n=1}^{n} \frac{1}{\lambda_n^2} - \cdots + (-1)^w \frac{z^w - z_0^w}{w} \sum_{n=1}^{n} \frac{1}{\lambda_n^w}.$$

Then

$$\Delta b_n(z) = b_n(z) [e^{p_{3n}} - 1],$$

where

$$p_{3n} = -(z - z_0) \frac{1}{\lambda_{n+1}} + \frac{z^2 - z_0^2}{2\lambda_{n+1}^2} - \cdots + (-1)^w \frac{z^2 - z_0^2}{w\lambda_{n+1}^w}.$$

Whence

$$L_n^{(1)}(z) = b_n(z) a_n \left[-(z - z_0) + \theta_1^{(1)} \frac{1}{\lambda_n} + \cdots + \theta_{r-1}^{(1)} \frac{1}{\lambda_n^r} + \theta_r^{(1)}(n) \frac{1}{\lambda_n^r} \right],$$

where $\theta_1^{(1)}, \cdots, \theta_{r-1}^{(1)}$ are independent of n and bounded in z and $\theta_r^{(1)}(n)$ is bounded in n and z. Apply the rule for differencing a product. In case λ_{n+j} should occur at any time instead of λ_n, replace it by

$$\lambda_n + \frac{k_1}{\lambda_n} + \frac{k_2}{\lambda_n^2} + \cdots + \frac{k_r(n)}{\lambda_n^r}$$

where $k_1, \cdots, k_{r-1}$ are constants and $k_r(n)$ is bounded. We obtain

$$L_n^{(k)}(z) = b_n(z)\,[(-1)^k \sigma_n^k (z - z_0)^k]\,[1 + \eta(n, z)] \qquad (k = 1, \cdots, r), \tag{9}$$

where $\eta(n, z) \to 0$ uniformly in z. In a similar way we show that

$$M_n^{(r)}(z) = \frac{1}{\lambda_n}\,[(-1)^{r+1}\sigma_n^r (z - z_0)^{r+1}]\,[1 + \eta'(n, z)],$$

where $\eta'(n, z) \to 0$ uniformly in z. Let $z = x + yi$. Now

$$\frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n^2} \to 0\,.$$

We, consequently, conclude that

$$\frac{|\, b_n(z) \,|}{b_n(x)} < M_1$$

independent of z and n. It results that

$$\frac{|\, M_n^{(r)}(z) \,|}{(-1)^{r+1} M_n^{(r)}(x)} < M \tag{10}$$

independent of z and n.

We are now in a position to apply Theorem III. From (8) and (9) hypotheses (ii) are seen to be fulfilled. By means of (10) we can prove the series in hypothesis (iii) convergent.

$$|\, M_n^{(r)}(z) \,| < (-1)^{r+1} M M_n^{(r)}(x). \tag{11}$$

To prove the series whose general term is the right member of (11) convergent, remark first that it is summable $R[r, \lambda]$. We illustrate with a first order sum. From (5)

$$\frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{1}{\lambda_n} \sum_{n=1}^{n} M_n^{(1)} s_n^{(1)}(x) = S_n^{(1)}(x) - b_{n+1}(x) s_n^{(1)}(x)$$
$$+ \frac{1}{\sigma_n} \sum_{n=1}^{n} [\Delta b_{n+1}(x)] \sigma_n s_n^{(1)}(x) + \frac{1}{\sigma_n} \sum_{n=1}^{n} \frac{\lambda_{n+1}}{\lambda_n} [\Delta b_{n+1}(x)]\, \sigma_n s_n^{(1)}(x).$$

In this let $s_n^{(1)}(x) \equiv 1$. That is, consider as a-series the series $1 + 0 + 0 + \cdots$. Thereupon in this case $S_n^{(1)}(x) = b_1(x)$ and

$$
\begin{aligned}
\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\sum_{n=1}^{n} M_n^{(1)}(x) &= b_1(x) - b_{n+2}(x) \\
&\quad + \frac{1}{\sigma_n}\sum_{n=1}^{n}[\Delta b_{n+1}(x)]\,\sigma_n + \frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{\lambda_{n+1}}{\lambda_n}[\Delta b_{n+1}(x)]\,\sigma_n .
\end{aligned} \tag{12}
$$

The right member approaches a limit when $n \to \infty$. Take for example one of the sums

$$
\frac{1}{\sigma_n}\sum_{n=1}^{n}\frac{1}{\lambda_n}\lambda_n\sigma_n\Delta b_{n+1}(x). \tag{13}
$$

We know by (9) that $\lambda_n\sigma_n\Delta b_{n+1}(x)$ approaches a limit. Consequently, (13) approaches a limit. This being true for all terms in the right member of (12), it is also true of the left member, that is,

$$
\sum_{n=1}^{\infty}\sigma_n\Delta\lambda_n\Delta b_n(x)
$$

is summable $R[1, \lambda]$. But by (9) the terms are all of the same sign when n is sufficiently great. Consequently, by Theorem C the series is convergent. Hence by (11)

$$
\sum_{n=1}^{\infty} M_n^{(1)}(z)s_n^{(1)}(z)
$$

is absolutely convergent. Combine this result with proved result that hypotheses (ii) of Theorem III are satisfied and we see that all hypotheses of Theorem III are fulfilled. The reasoning in the general case is identical with that just given.

3. **Factorial series.** Consider the factorial series

$$
\sum_{n=1}^{\infty} C_n\frac{\lambda_1\cdots\lambda_n}{(z+\lambda_1)\cdots(z+\lambda_n)}, \tag{14}
$$

where λ_n is subject to the same restrictions as in the previous section plus the restrictions that $(\sigma_n^r/\lambda_n^w) \to$ a limit and that $\sum_{n=1}^{\infty}(\sigma_n^r/\lambda_n^{w+1})$ be convergent.

We shall prove the following theorem.

THEOREM V. *(14) is summable $R|r, \lambda|$ at the same points as (7) with the exception of $-\lambda_1, -\lambda_2, \cdots$. It is uniformly summable $R|r, \lambda|$ over any bounded region P over which (7) is uniformly summable $R|r, \lambda|$ but from which small circular regions about $-\lambda_1, -\lambda_2, \cdots$ have been deleted.*

To prove this theorem we begin by writing (14) in the following form

$$\sum_{n=1}^{\infty} C_n e^{p_{1n}} b_n ,$$

where

$$b_n = \frac{\lambda_1 \cdots \lambda_n}{(z + \lambda_1) \cdots (z + \lambda_n)} e^{-p_{1n}}.$$

Throughout the proof we shall assume $\lambda_1 > |z| + \epsilon$. In this there is no loss of generality since $0 < \lambda_n \to \infty$ and since the mere fact of R-summability is not affected by the omission of a fixed number of terms at the beginning of a series nor by multiplying through by a function independent of n. We shall find that all conditions imposed by Theorem III on b_n are fulfilled. We readily show that $b_n = e^{p_{4n}}$, where

$$p_{4n} = (-1)^{w+1} \frac{z^{w+1}}{w+1} \sum_{n=1}^{n} \frac{1}{\lambda_n^{w+1}} + \cdots$$

$$+ (-1)^{w+r} \frac{z^{w+r}}{w+r} \sum_{n=1}^{n} \frac{1}{\lambda_n^{w+r}} + \sum_{n=1}^{n} \theta\left(\frac{z}{\lambda_n}\right) \frac{1}{\lambda_n^{w+r+1}},$$

where $\theta(z/\lambda_n)$ is bounded.

$$\Delta b_n = b_n(e^{p_{5n}} - 1),$$

where

$$p_{5n} = (-1)^{w+1} \frac{z^{w+1}}{\lambda_{n+1}^{w+1}} + \cdots + \theta\left(\frac{z}{\lambda_n}\right) \frac{1}{\lambda_{n+1}^{w+r+1}}.$$

From this

$$\Delta\sigma_n \lambda_n \Delta b_n(z) = b_n \sigma_n \Bigg[(-1)^{w+1} \frac{z^{w+1}}{\lambda_n^{w+1}} + h_2^{(1)}(z) \frac{1}{\lambda_n^{w+2}}$$

$$+ \cdots + h_{r-1}^{(1)}(z) \frac{1}{\lambda_n^{w+r-1}} + H^{(1)}(z, \lambda_n) \frac{1}{\lambda_n^{w+r}} \Bigg],$$

where $h_j^{(1)}(z)$ is independent of n, $j = 2, \cdots, r - 1$ and $H^{(1)}$ is bounded. Continuing, we have the result that

$$L_n^{(r)}(z) = b_n \sigma_n^r H^{(r)} \frac{1}{\lambda_n^w},$$

where $H^{(r)}$ is bounded. Similarly

$$M_n^{(r)}(z) = b_n \sigma_n^r \bar{H}^{(r)} \frac{1}{\lambda_n^{w+1}},$$

where $\bar{H}^{(r)}$ is bounded. It now is immediate that b_n fulfills the conditions imposed by Theorem III.

We now turn about and consider series (7). We let

$$b_n = e^{p_{1n}} \frac{(z + \lambda_1) \cdots (z + \lambda_n)}{\lambda_1 \cdots \lambda_n}.$$

The minus sign is immaterial to the argument already given on b_n. Consequently, b_n in this case satisfies all the conditions imposed by Theorem III and the proof of Theorem V is complete.

Theorem V combined with Theorem IV establish a half-plane as the region of summability of (14).

LEHIGH UNIVERSITY.

ions

onse-
III

egion